SEQUENCE LISTING

```
<110> WAKITA, Takaji
      KATO, Takanobu
      DATE, Tomoko
<120> A NUCLEIC ACID CONSTRUCT CONTAINING A NUCLEIC ACID DERIVED
      FROM THE GENOME OF HEPATITIS C VIRUS (HCV) OF GENOTYPE 2a,
     AND A CELL HAVING SUCH NUCLEIC ACID CONSTRUCT INTRODUCED THEREIN
<130> 1254-0299PUS1
<140> US 10/558,155
<141> 2005-11-23
<150> PCT/JP2003/015038
<151> 2003-11-25
<150> JP 2003-148242
<151> 2003-05-26
<150> JP 2003-329115
<151> 2003-09-19
<160> 41
<170> PatentIn Ver. 2.1
<210> 1
<211> 8024
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: replicon
<400> 1
accugecceu aauaggggeg acacueegee augaaucaeu eeecugugag gaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
ccccucccq qqaqaqccau aquqqucuqc qgaaccqquq aquacaccqq aauuqccqqq 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caaqacuqcu agccqaquag cguuqgguug cgaaaggccu ugugguacug ccugauaggg 300
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa auccuaaacc 360
ucaaaqaaaa accaaaaqaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420
cqcaqquucu ccqqccqcuu qqquqqaqaq gcuauucggc uaugacuggg cacaacagac 480
aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggcgcc cgguucuuuu 540
uqucaaqacc qaccuquccq quqcccuqaa uqaacuqcaq qacqaqqcaq cqcqqcuauc 600
quqqcuqqcc acqacqqqcq uuccuuqcqc aqcuquqcuc qacquuquca cuqaaqcqqq 660
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720
uccugccgag aaaguaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780
qqcuaccuqc ccauucqacc accaaqcqaa acaucqcauc qaqcqaqcac quacucqgau 840
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900
```

cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960 uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020 cuguggccgg cuggggugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080 ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140

ucccgauucg	cagcgcaucg	ccuucuaucg	ccuucuugac	gaguucuucu	gaguuuaaac	1200
ccucucccuc	cccccccu	aacguuacug	gccgaagccg	cuuggaauaa	ggccggugug	1260
cguuugucua	uauguuauuu	uccaccauau	ugccgucuuu	uggcaaugug	agggcccgga	1320
aaccuggccc	ugucuucuug	acgagcauuc	cuaggggucu	uuccccucuc	gccaaaggaa	1380
	guugaauguc				ugaagacaaa	1440
	agcgacccuu					1500
	gccacgugua					1560
	gauaguugug					1620
	ugcccagaag					1680
	auguguuuag					1740
	ccuuugaaaa					1800
	gccuccuggg					1860
	aaguccaaau					1920
	uguggacugu					1980
					gcccagcccc	2040
					ucuggucacg	
cggaacgcug	augucauccc	ggcucggaga	cgcggggaca	agcggggagc	auugcucucc	2160
ccgagaccca	uuucgaccuu	gaaggggucc	ucgggggggc	cggugcucug	cccuaggggc	2220
cacqucguug	ggcucuuccg	agcagcugug	ugcucucggg	gcguggccaa	auccaucgau	2280
uucauccccg	uuqagacacu	cgacguuguu	acaaggucuc	ccacuuucag	ugacaacagc	2340
acgccaccgg	cugugcccca	gaccuaucag	gucggguacu	ugcaugcucc	aacuggcagu	2400
ggaaagagca	ccaagguccc	ugucgcguau	gccgcccagg	gguacaaagu	acuagugcuu	2460
aaccccucgg	uagcugccac	ccugggguuu	ggggcguacc	uauccaaggc	acauggcauc	2520
aaucccaaca	uuaggacugg	agucaggacc	gugaugaccg	gggaggccau	cacguacucc	2580
acauauggca	aauuucucgc	cgaugggggc	ugcgcuagcg	gcgccuauga	caucaucaua	2640
ugcqauqaau	gccacgcugu	ggaugcuacc	uccauucucg	gcaucggaac	gguccuugau	2700
caaqcaqaqa	cagccggggu	cagacuaacu	gugcuggcua	cggccacacc	ccccggguca	2760
gugacaaccc	cccaucccga	uauagaagag	guaggccucg	ggcgggaggg	ugagaucccc	2820
uucuauggga	gggcgauucc	ccuauccugc	aucaagggag	ggagacaccu	gauuuucugc	2880
cacucaaaqa	aaaaguguga	cgagcucgcg	gcggcccuuc	ggggcauggg	cuugaaugcc	2940
	auagaggguu					3000
	acgcccucau					3060
	ucacccaagc					3120
	cacaagacgc					3180
	cuuauaggua					3240
	gugagugcua					3300
accaccquca	ggcuuagagc	guauuucaac	acgcccggcc	uacccgugug	ucaagaccau	3360
cuuqaauuuu	gggaggcagu	uuucaccggc	cucacacaca	uagacgccca	cuuccucucc	3420
caaacaaagc	aagcggggga	gaacuucgcg	uaccuaguag	ccuaccaagc	uacggugugc	3480
gccagagcca	aggccccucc	cccquccugg	gacgccaugu	ggaagugccu	ggcccgacuc	3540
					uaccaaugag	3600
	cacacccugg					3660
	gcacgugggu					3720
cuggcgacug	gaugcguuuc	caucaucggc	cgcuugcacg	ucaaccagcg	agucgucguu	3780
gcgccggaua	aggagguccu	guaugaggcu	uuugaugaga	uggaggaaug	cgccucuagg	3840
geggeueuca	ucqaaqaqqq	gcagcggaua	gccgagaugu	ugaaguccaa	gauccaaggc	3900
uuqcuqcaqc	aggccucuaa	gcaggcccag	gacauacaac	ccgcuaugca	ggcuucaugg	3960
cccaaaquqq	aacaauuuug	ggccagacac	auguggaacu	ucauuagcgg	cauccaauac	4020
cucqcaqqau	uqucaacacu	gccagggaac	cccgcggugg	cuuccaugau	ggcauucagu	4080
geegeeeuca	ccaquccquu	gucgaccaqu	accaccaucc	uucucaacau	caugggaggc	4140
					caguggccug	4200
quaqaaacua	ccquqqqcaq	cauaggccug	gguaaggugc	ugguggacau	ccuggcagga	4260
uauggugcag	gcauuucqaa	ggcccucquc	gcauucaaga	ucaugucuqq	cgagaagccc	4320
ucuauggaag	augucaucaa	ucuacugccu	gggauccugu	cuccgggagc	ccugguggug	4380
ggggucaucu	gcgcggccau	ucugegeege	cacguaggac	cgggggaggg	cgcgguccaa	4440
uggaugaaca	ggcuuauugc	cuuugcuucc	agaggaaacc	acgucgcccc	uacucacuac	4500
gugacggagu	cggaugcguc	gcagcgugug	acccaacuac	uuggcucucu	uacuauaacc	4560

	gaagacucca					4620
	gcgacgugug					4680
	aauuguuccc					4740
	ugugggccgg					4800
	auguccgccu					4860
	ggaccuuucc					4920
	acaagaccgc					4980
	cguacuccua					5040
	cuccagaguu					5100
	agccguuuuu					5160
	cccagcuucc					5220
	cgccccacau					5280
	aggcgagcuc					5340
	acagcaacac					5400
	cucagacaga					5460
	aagagagcga					5520
	cacgggccuu					5580
	ggaggccaga					5640
	ccccgacgcc					5700
	aagcccucca					5760
	gcucguccac					5820
	ccccucaga					5880
	cggaccugga					5940
	ccgguucggg					6000
	gcuccauguc					6060
	aguugccaau					6120
	caacaucaaa					6180
	ucgacgccca					6240
	caaggcuccu					6300
	aguauggauu					6360
	aguccgugug					6420
	ccaaaaauga					6480
	ucguuuaccc					6540 6600
gacauuacac	aaaagcuucc	ucaggcggua	augggagcuu	ccuauggcuu	ccaguacucc	
ccugcccaac	ggguggagua	ucucuugaaa	gcaugggcgg	aaaagaagga	ccccaugggu	6720
	auacccgaug					6780
	accaggccug					6840
	uuuacguagg					
					cacaugcuau	6960
	uagcggccug					7020
	uaguagucau					7080
	cggaggccau					7140
	uggagcuaau					7200
	gcagauacua					7260
ccaaccay	uuagacacuc	adilconssila	acauggeugg	ucuccauucu	caugguccaa	
ascaccauau	accadaaccu	caacuuugag	acacacaaca	cacuauacuc	cgugaauccu	7380
yacacccuyg	caccagaaccu	ugagagguua	cacadaciiia	acaccininic	uaugcacaca	7440
nacucucacc	acqaaciiqac	acadanaach	ucadeceiica	dagaaciiida	ggcgccaccc	7500
	ggaagagucg					7560
	uuugcggccg					7620
	cggaggcgcg					7680
	uuuuucacag					7740
					cggcacacac	
naddnacach	ccanadenaa	channeemin	111111111111111111111111111111111111111		uuuuuuuuu	7860
uayyaacaca	Chinnin	nnnnccene	uuucuucccu	ucucaucuua	uucuacuuuc	7920
					cgugagccgc	
	J = = = = = = = = = = = = = = = = = = =	J	J J J	J J 99	J J J 9 -	-

```
<210> 2
<211> 8024
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: replicon
<400> 2
acccgccccu aauaggggcg acacuccgcc augaaucacu ccccugugag gaacuacugu 60
cuucacqcaq aaaqcqucua qccauggcgu uaguaugagu gucguacagc cuccaggccc 120
ccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
ugcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa aucccaaacc 360
ucaaagaaaa accaaaagaa acacuaaccg ucgcccaaug auugaacaag auggauugca 420
cgcagguucu ccggccgcuu ggguggagag gcuauucggc uaugacuggg cacaacagac 480
aaucggcugc ucugaugccg ccguguuccg gcugucagcg caggggcgcc cgguucuuuu 540
ugucaagacc gaccuguccg gugcccugaa ugaacugcag gacgaggcag cgcggcuauc 600
guggcuggcc acgacgggcg uuccuugcgc agcugugcuc gacguuguca cugaagcggg 660
aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720
uccuqccqaq aaaquaucca ucauggcuga ugcaaugcgg cggcugcaua cgcuugaucc 780
ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840
ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccagc 900
cgaacuguuc gccaggcuca aggcgcgcau gcccgacggc gaggaucucg ucgugaccca 960
uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020
cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080
ugcuqaaqaq cuuqqcqqcq aauqqqcuqa ccqcuuccuc quqcuuuacq guaucqccqc 1140
uccegauucg cagegeaucg ceuucuaucg ceuucuugac gaguucuucu gaguuuaaac 1200
ccucucccuc cccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260
cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320
aaccuggeee ugucuucuug aegageauue euaggggueu uueeeeueue geeaaaggaa 1380
ugcaaggucu guugaauguc gugaaggaag caguuccucu ggaagcuucu ugaagacaaa 1440
caacgucugu agcgacccuu ugcaggcagc ggaacccccc accuggcgac aggugccucu 1500
gcggccaaaa gccacgugua uaagauacac cugcaaaggc ggcacaaccc cagugccacg 1560
uugugaguug gauaguugug gaaagaguca aauggcucuc cucaagcgua uucaacaagg 1620
ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680
caugcuuuac auguguuuag ucgagguuaa aaaaacgucu aggccccccg aaccacgggg 1740
acgugguuuu ccuuugaaaa acacgauaau accauggccc ccaucaccgc uuacgcccag 1800
cagacacgag gucucuuggg cucuauagug gugagcauga cgggggguga caagacagaa 1860
caggccgggg agguccaagu ccuguccaca gucacucagu ccuuccucgg aacauccauu 1920
ucqqqqqucu uauqqacugu uuaccacgga gcuggcaaca agacacuagc cggcucgcgg 1980
ggcccgguca cgcagaugua cucgagcgcc gagggggacu uggucgggug gcccagcccu 2040
ccugggacca aaucuuugga geeguquacq uguggagegg ucgaccugua uuuggucacg 2100
cggaacgcug augucauccc ggcucgaaga cgcggggaca agcggggagc gcugcucucc 2160
ccqaqacccc uuucgaccuu gaaggggucc ucggggggac cugugcuuug cccuaggggc 2220
cacgcugucg gaaucuuccg ggcagcugug ugcucucggg guguggcuaa guccauagau 2280
uucauccccg uugagacgcu cgacaucguc acgcggucuc ccaccuuuag ugacaacagc 2340
acaccaccag cugugcccca gaccuaucag gugggguacu ugcacgcccc cacuggcagu 2400
ggaaaaagca ccaagguccc cgucgcguac gccgcccagg gguauaaagu gcuggugcuc 2460
aaucccucqq uqqcuqccac ccuqqqauuu qqqqcquacu uguccaaggc acauggcauc 2520
aaccccaaca uuaggacugg agucagaacu gugacgaccg gggagcccau uacauacucc 2580
acguauggua aauuccucgc cgaugggggc ugcgcaggcg gcgccuauga caucaucaua 2640
ugcgaugaau gccacucugu ggaugcuacc acuauucucg gcaucgggac aguccuugac 2700
```

caagcagaga	cagccggggu	caggcuaacu	quacuggcca	cadccacacc	ccccgggucg	2760
quqacaaccc	cccaucccaa	uauagaggag	guageceueg	gacaggaggg	ugagaucccc	2820
					gauuuucugc	
cacucaaaga	aaaaguguga	cgagcucgca	acggcccuuc	ggggcauggg	cuugaacgcu	
anaacanann	acagaggguu	adacanence	auaauaccaa	cucaaggaga	uanaanaanc	3000
					gaucgacugc	
aacdnadcdd	ncacccaddc	canadacinc	addenidaece	ccaccuucac	uauaaccaca	3120
cadacuducc	cacaagacac	nancheacan	agucaggacc	gagggggac	ggguagagga	3180
adachadaca	uuuauaaaua	nanninceach	adnagacasa	cciicaddaaii	guuugacagu	3240
quaquacucu	gugagugcua	cdacacada	achachhaan	andadenene	accadnadad	3300
accaccouca	addicadaa	qualiturcaac	acaccuaacu	nacchanana	ccaggaccac	
cuidadiiiii	ggcacagggc	uuucaccaac	cucacacaca	nadacachca	uuuccuuucc	3420
cauguguuuu	aggaggeaga	aaauuucaca	nachhadhad	ccualicaddc	cacagugugc	3480
	aagcgcccc					3540
						3600
	uugugggccc					3660
	cacaccccgu					3720
	gcacgugggu				, ,	3780
	gguguguuuc					3840
	aggagguccu					3900
	uugaagaggg					3960
	aagccucuaa					4020
	agcaauucug					
					ggcauucagc	4140
	ccaguccguu					
	cccaaauugc					4200 4260
	cuguuggcag					4320
	gcauuucggg					
	augucaucaa					4380
	gcgcggccau					4440
	ggcuuaucgc					4500
					cacuauaacu	4560
	ggagacuuca					4620
	gcgaugugug					4680
	agcuguuccc					4740
	ugugggccgg					4800
	acguccgcuu					4860
	ggaccuuucc					4920
	ucaagaccgc					4980
	cauaugccua				,	5040
caacuccccu	cuccagaguu	uuucucuugg	guggacggag	uacaaaucca	uagguccgcc	5100
cccacaccaa	agccguuuuu	ccgggaugag	gucucguuca	gcguugggcu	caauucauuu	5160
gucgucgggu	cucagcuucc	cugugacccu	gagcccgaca	cugagguagu	gauguccaug	5220
cuaacagacc	caucccauau	cacggcggag	gcugcagcgc	ggcguuuagc	gcggggguca	5280
ccccaucug	aggcaagcuc	cucagcgagc	cagcugucgg	cgccaucgcu	gcgagccacc	5340
ugcaccaccc	acgguaggac	cuaugaugug	gacauggugg	augccaaccu	guucaugggg	5400
					ccucgacuca	5460
augaccgagg	aagagggcga	ccuugagccu	ucaguaccau	cggaguauau	gcuccccagg	5520
aagagguucc	caccggccuu	accggcuugg	gcgcggccug	auuacaaccc	accgcuugug	5580
gaaucgugga	agaggccaga	uuaccaacca	cccacuguug	cgggcugugc	ucucccccc	5640
cccaaaaaga	ccccgacgcc	uccuccaagg	agacgccgga	cagugggucu	gagcgagagc	5700
accauaggag	augcccucca	acagcuggcc	aucaaguccu	uuggccagcc	cccccaagc	5760
ggcgauucag	gccuuuccac	gggggcggac	gccgccgacu	ccggcgaucg	gacacccccu	5820
gacgaguugg	cucuuucgga	gacagguucu	accuccucca	ugccccccu	cgagggggag	5880
	cagaccugga					5940
					ggaugacucc	
gucgugugcu	gcuccauguc	auauuccugg	accggggcuc	uaauaacucc	uuguagcccc	6060
gaagaggaaa	aguugccaau	uaacuccuug	agcaacucgc	uguugcgaua	ccauaacaag	6120

```
augcaagugc ucgacgccua uuaugauuca gucuuaaagg acaucaagcu agcggccucc 6240
aaggucagcg caaggcuccu caccuuagag gaggcgugcc aauugacccc accccacucu 6300
gcaagaucca aguauggguu uggggcuaag gagguccgca gcuuguccgg gagggccguc 6360
aaccacauca aquecququq qaaqqaccuc uuqqaaqacu cacaaacacc aauuccuaca 6420
accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaaaaaacca 6480
gcucgccuua ucguuuaccc ugaccucggc gucagggucu gcgagaagau ggcccuuuau 6540
gaugucacac aaaagcuucc ucaggcggug augggggcuu cuuauggcuu ccaguacucc 6600
cccgcucagc ggguggaguu ucucuugaag gcaugggcgg aaaagagaga cccuaugggu 6660
uuuucguaug auacccgaug cuuugacuca accgucacug agagagacau caggacugag 6720
gaguccauau accaggecug cuccuuacce gaggaggece gaacugecau acaeuegeug 6780
acugagagac ucuauguggg agggcccaug uucaacagca agggccaguc cugcggguac 6840
aggeguugee gegeeagegg ggugeuuaee acuaguaugg ggaacaeeau cacaugeuau 6900
guaaaagccc uagcggcuug caaggcugcg gggauaauug cgcccacgau gcugguaugc 6960
ggcgacgacu uggucgucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020
agagecuuca eggaggeuau gaceagguau ueugeeceue euggugaeee eeceagaeeg 7080
gaauaugacc uggagcuaau aacaucuugu uccucaaacg ugucuguggc acuuggccca 7140
cagggccgcc gcagauacua ccugaccaga gaccccacca cuucaauugc ccgggcugcc 7200
ugggaaacag uuagacacuc cccugucaau ucauggcugg gaaacaucau ccaguacgcu 7260
ccaaccauau ggguucgcau gguccugaug acacacuucu ucuccauucu cauggcccag 7320
gacacccuag accagaaccu uaacuuugaa auguacggau cgguguacuc cgugaguccu 7380
cuggaccucc cagccauaau ugaaagguua cacgggcuug acgccuucuc ucugcacaca 7440
uacacuccec acgaacugac geggguggeu ucageceuca gaaaacuugg ggegeeacec 7500
cucagagegu ggaagagueg ggegegugea guuagggegu eeeucaueue eegugggggg 7560
agggcggccg uuugcggucg guaccucuuc aacugggcgg ugaagaccaa gcucaaacuc 7620
acuccuuugc cggaggcacg ccuccuggau uuguccaguu gguuuaccgu cggcgccggc 7680
gggggcgaca uuuaucacag cgugucgcgu gcccgacccc gccuauuacu ccuuagccua 7740
cuccuacuuu cuguaggggu aggccucuuc cuacuccccg cucgauagag cggcacacau 7800
uuuuuuuuu cuuuuuuuu uuuuucccuc uuucuucccu ucucaucuua uucuacuuuc 7920
uuucuuggug gcuccaucuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980
                                                                 8024
augacugcag agagugccgu aacuggucuc ucugcagauc augu
<210> 3
<211> 9678
<212> DNA
<213> Hepatitis C virus
<220>
<221> CDS
<222> (341)..(9442)
<400> 3
acctgcccct aataggggcg acactccgcc atgaatcact cccctgtgag gaactactgt 60
cttcacgcag aaagcgccta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc 120
ccccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg 180
aagactgggt cetttettgg ataaacceae tetatgeeeg gecatttggg egtgeeeeeg 240
caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg 300
                                                                 355
cgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat cct
                                           Met Ser Thr Asn Pro
```

guauacugua cuacaucaaa gagugccuca cuaagggcua aaaagguaac uuuugauagg 6180

	caa Gln							403
	ttc Phe							451
	agg Arg 40							499
	tcc Ser							547
	act Thr							595
	aat Asn	 		 -	 	_		643
	cgc Arg							691
	ggt Gly 120							739
	tac Tyr							787
	gcg Ala							835
	aac Asn							883
	tgc Cys							931
	agc Ser 200							979
 _	ctc Leu	 -	 -					1027

					acg Thr 235											1075
					ccc Pro											1123
					atg Met											1171
					ggg											1219
_	_				tgg Trp											1267
					gga Gly 315											1315
					acc Thr											1363
					atc Ile											1411
					tct Ser											1459
		_	_	_	gct Ala			-								1507
					tcc Ser 395											1555
					att Ile											1603
					ttg Leu											1651
					tac Tyr											1699
ggg	cgc	ctg	tcc	gcc	tgc	cgc	aac	atc	gag	gct	ttc	cgg	ata	ggg	tgg	1747

Gly	Arg 455	Leu	Ser	Ala	Cys	Arg 460	Asn	Ile	Glu	Ala	Phe 465	Arg	Ile	Gly	Trp	
											cca Pro					1795
											ggc Gly					1843
											ccc Pro					1891
											tac Tyr					1939
											cga Arg 545					1987
											ggt Gly					2035
											ttc Phe					2083
											cat His					2131
		_	_								cca Pro	-	_	_	_	2179
cac His	tac Tyr 615	cct Pro	tac Tyr	aga Arg	ctc Leu	tgg Trp 620	cat His	tac Tyr	ccc Pro	tgc Cys	aca Thr 625	gtc Val	aat Asn	ttt Phe	acc Thr	2227
											gag Glu					2275
											gac Asp					2323
											acc Thr					2371
											ttg Leu					2419

		000					003					000				
				cag Gln												2467
tca Ser 710	cct Pro	gct Ala	atc Ile	aca Thr	aaa Lys 715	tac Tyr	gtc Val	gtt Val	cga Arg	tgg Trp 720	gag Glu	tgg Trp	gtg Val	gta Val	ctc Leu 725	2515
				tta Leu 730												2563
				ggc Gly												2611
				agt Ser												2659
				gca Ala												2707
				ctc Leu												2755
				cag Gln 810												2803
				ttg Leu												2851
ggg Gly	tat Tyr	aag Lys 840	acc Thr	ctc Leu	ctc Leu	Gly	cag Gln 845	Cys	ctg Leu	tgg Trp	Trp	ttg Leu 850	Cys	tat Tyr	ctc Leu	2899
				gaa Glu												2947
				cgc Arg												2995
				ttt Phe 890												3043
				tta Leu												3091

			gct ttg gtg a Ala Leu Val 1		
			ttg gcc ctt of Leu Ala Leu (945		
			cct atg tcg of Pro Met Ser i 960		
			gtg gaa ccc a Val Glu Pro 1 975		
Pro Met Glu			gga gcg gag a Gly Ala Glu '		
			tcc gcc cga (Ser Ala Arg 1		
	Gly Pro Ala		acc tcc aag of Thr Ser Lys of 1025		
-	_	-	caa aca cga d Gln Thr Arg d 1040	Gly Leu Leu	
		Thr Gly Arg	gac agg aca Asp Arg Thr 1055		
Ğlu Val Gln			cag tcc ttc Gln Ser Phe		
			cac gga gct His Gly Ala		
	Leu Arg Gly		cag atg tac Gln Met Tyr 1105		
			cct ggg acc Pro Gly Thr 1120	Lys Ser Leu	
		Val Asp Leu	tat ctg gtc Tyr Leu Val 1135		

gat gtc atc ccg gct cgg aga cgc ggg gac aag cgg gga gca ttg ctc Asp Val Ile Pro Ala Arg Arg Arg Gly Asp Lys Arg Gly Ala Leu Leu 1145 1150 1155	3811
tcc ccg aga ccc att tcg acc ttg aag ggg tcc tcg ggg ggg ccg gtg Ser Pro Arg Pro Ile Ser Thr Leu Lys Gly Ser Ser Gly Gly Pro Val 1160 1165 1170	3859
ctc tgc cct agg ggc cac gtc gtt ggg ctc ttc cga gca gct gtg tgc Leu Cys Pro Arg Gly His Val Val Gly Leu Phe Arg Ala Ala Val Cys 1175 1180 1185	3907
tct cgg ggc gtg gcc aaa tcc atc gat ttc atc ccc gtt gag aca ctc Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile Pro Val Glu Thr Leu 1190 1195 1200 1205	3955
gac gtt gtt aca agg tct ccc act ttc agt gac aac agc acg cca ccg Asp Val Val Thr Arg Ser Pro Thr Phe Ser Asp Asn Ser Thr Pro Pro 1210 1215 1220	4003
gct gtg ccc cag acc tat cag gtc ggg tac ttg cat gct cca act ggc Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu His Ala Pro Thr Gly 1225 1230 1235	4051
agt gga aag agc acc aag gtc cct gtc gcg tat gcc gcc cag ggg tac Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr Ala Ala Gln Gly Tyr 1240 1245 1250	4099
aaa gta cta gtg ctt aac ccc tcg gta gct gcc acc ctg ggg ttt ggg Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala Thr Leu Gly Phe Gly 1255 1260 1265	4147
gcg tac cta tcc aag gca cat ggc atc aat ccc aac att agg act gga Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro Asn Ile Arg Thr Gly 1270 1275 1280 1285	4195
gtc agg acc gtg atg acc ggg gag gcc atc acg tac tcc aca tat ggc Val Arg Thr Val Met Thr Gly Glu Ala Ile Thr Tyr Ser Thr Tyr Gly 1290 1295 1300	4243
aaa ttt ctc gcc gat ggg ggc tgc gct agc ggc gcc tat gac atc atc Lys Phe Leu Ala Asp Gly Gly Cys Ala Ser Gly Ala Tyr Asp Ile Ile 1305 1310 1315	4291
ata tgc gat gaa tgc cac gct gtg gat gct acc tcc att ctc ggc atc Ile Cys Asp Glu Cys His Ala Val Asp Ala Thr Ser Ile Leu Gly Ile 1320 1325 1330	4339
gga acg gtc ctt gat caa gca gag aca gcc ggg gtc aga cta act gtg Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly Val Arg Leu Thr Val 1335 1340 1345	4387
ctg gct acg gcc aca ccc ccc ggg tca gtg aca acc ccc cat ccc gat Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr Thr Pro His Pro Asp 1350 1355 1360 1365	4435
ata gaa gag gta ggc ctc ggg cgg gag ggt gag atc ccc ttc tat ggg	4483

Ile	Glu	Glu		Gly 1370	Leu	Gly	Arg		Gly .375	Glu	Ile	Pro		Tyr .380	Gly	
		Ile		cta Leu			Ile					His				4531
	His			aaa Lys		Cys					Ala					4579
Met				gcc Ala	Val					Gly						4627
	Pro			gga Gly					Val					Leu		4675
			Thr	gga Gly 1450				Ser					Asn			4723
		Gln		gtc Val			Ser					Phe				4771
	Gln			cca Pro		Asp					Ser					4819
Arg				gga Gly	Arg					Arg						4867
	Arg			gga Gly					Val					Cys		4915
			Ala	gcg Ala 1530				Leu					Thr			4963
		Arg		tat Tyr			Thr					Val				5011
	Leu			tgg Trp		Ala					Leu					5059
Ala				tcc Ser	Gln					Gly						5107
				caa Gln												5155

1590	1595	1600	1605
Pro Ser Trp Asp		tgc ctg gcc cga ct Cys Leu Ala Arg Le 1615	
	Thr Pro Leu Leu	tac cgt ttg ggc co Tyr Arg Leu Gly P: 1630	
		acg aag tac atc go Thr Lys Tyr Ile A	la Thr Cys Met
		agc acg tgg gtc c Ser Thr Trp Val Lo 1665	
		tgc ctg gcg act g Cys Leu Ala Thr G 1680	
Ile Ile Gly Arg		cag cga gtc gtc g Gln Arg Val Val V 1695	
	Tyr Glu Ala Phe	gat gag atg gag g Asp Glu Met Glu G 1710	
		cag cgg ata gcc g Gln Arg Ile Ala G 17	lu Met Leu Lys
		cag gcc tct aag c Gln Ala Ser Lys G 1745	
		tgg ccc aaa gtg g Trp Pro Lys Val G 1760	
Ala Arg His Met		agc ggc atc caa t Ser Gly Ile Gln T 1775	
		gcg gtg gct tcc a Ala Val Ala Ser M 1790	
		tcg acc agt acc a Ser Thr Ser Thr T 18	hr Ile Leu Leu
		tcc cag atc gca c Ser Gln Ile Ala P 1825	

gcc acc ggc ttt o Ala Thr Gly Phe V 1830		y Ala Ala Val Gl	
ata ggc ctg ggt a Ile Gly Leu Gly 1			y Ala
ggc att tcg ggg of Gly Ile Ser Gly 1865			
ccc tct atg gaa o Pro Ser Met Glu 7 1880	 sn Leu Leu Pro		-
gga gcc ctg gtg o Gly Ala Leu Val ' 1895			
gtg gga ccg ggg Val Gly Pro Gly 1910		t Asn Arg Leu Il	
ttt gct tcc aga Phe Ala Ser Arg			nr Glu
tcg gat gcg tcg Ser Asp Ala Ser 1945			
acc agc cta ctc Thr Ser Leu Leu 1960	 is Asn Trp Ile		
atc cca tgc tcc Ile Pro Cys Ser 1975	 		
acc atc ttg aca Thr Ile Leu Thr 1990		r Ser Lys Leu Pl	
aag ctg ccc ggc Lys Leu Pro Gly 2			ys Gly
gtg tgg gcc ggc Val Trp Ala Gly 2025			
aac atc tct ggc Asn Ile Ser Gly 2040	eu Gly Ser Me		

aaa acc tgc atg aac acc Lys Thr Cys Met Asn Thr 2055	Trp Gln Gly Thr Phe	* * * · · · · · · · · · · · · · · · · ·	6547
acg gag ggc cag tgc gcg Thr Glu Gly Gln Cys Ala 2070 2075			6595
atc tgg agg gtg gcg gcc Ile Trp Arg Val Ala Ala 2090			6643
tcg tac tcc tat gta aca Ser Tyr Ser Tyr Val Thr 2105			6691
tgc caa cta cct tct cca Cys Gln Leu Pro Ser Pro 2120			6739
atc cat agg ttt gca ccc Ile His Arg Phe Ala Pro 2135	Thr Pro Lys Pro Phe		6787
tcg ttc tgc gtt ggg ctt Ser Phe Cys Val Gly Leu 2150 2155		9999	6835
tgt gaa cct gag ccc gac Cys Glu Pro Glu Pro Asp 2170			6883
ccg ccc cac atc acg gcg Pro Pro His Ile Thr Ala 2185		-9 9 9 9 9 9 9	6931
tca cct cca tct gag gcg Ser Pro Pro Ser Glu Ala 2200		049 004 04- 9 9	6979
tcg ctg cgg gcc acc tgc Ser Leu Arg Ala Thr Cys 2215	Thr Thr His Ser Asn	9 9-5	7027
atg gtc gat gcc aac ctg Met Val Asp Ala Asn Leu 2230 2235		5 + 5 - 5 5 - 5 - 5 - 5 - 5 - 5	7075
cct gag tcc agg gtg ccc Pro Glu Ser Arg Val Pro 2250			7123
gaa gag agc gac ctt gag Glu Glu Ser Asp Leu Glu 2265			7171

agg agc ggg ttt cca cgg gcc tta ccg gct tgg gca cgg cct gac tac Arg Ser Gly Phe Pro Arg Ala Leu Pro Ala Trp Ala Arg Pro Asp Tyr 2280 2285 2290	7219
aac ccg ccg ctc gtg gaa tcg tgg agg agg cca gat tac caa ccg ccc Asn Pro Pro Leu Val Glu Ser Trp Arg Arg Pro Asp Tyr Gln Pro Pro 2295 2300 2305	7267
acc gtt gct ggt tgt gct ctc ccc ccc ccc aag aag gcc ccg acg cct Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys Lys Ala Pro Thr Pro 2310 2315 2320 2325	7315
ccc cca agg aga cgc cgg aca gtg ggt ctg agc gag agc acc ata tca Pro Pro Arg Arg Arg Thr Val Gly Leu Ser Glu Ser Thr Ile Ser 2330 2335 2340	7363
gaa gcc ctc cag caa ctg gcc atc aag acc ttt ggc cag ccc ccc tcg Glu Ala Leu Gln Gln Leu Ala Ile Lys Thr Phe Gly Gln Pro Pro Ser 2345 2350 2355	7411
agc ggt gat gca ggc tcg tcc acg ggg gcg ggc gcc gcc gaa tcc ggc Ser Gly Asp Ala Gly Ser Ser Thr Gly Ala Gly Ala Ala Glu Ser Gly 2360 2365 2370	7459
ggt ccg acg tcc cct ggt gag ccg gcc ccc tca gag aca ggt tcc gcc Gly Pro Thr Ser Pro Gly Glu Pro Ala Pro Ser Glu Thr Gly Ser Ala 2375 2380 2385	7507
tcc tct atg ccc ccc ctc gag ggg gag cct gga gat ccg gac ctg gag Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly Asp Pro Asp Leu Glu 2390 2395 2400 2405	7555
tct gat cag gta gag ctt caa cct ccc ccc cag ggg ggg ggg gta gct Ser Asp Gln Val Glu Leu Gln Pro Pro Pro Gln Gly Gly Val Ala 2410 2415 2420	7603
ccc ggt tcg ggc tcg ggg tct tgg tct act tgc tcc gag gag gac gat Pro Gly Ser Gly Ser Gly Ser Trp Ser Thr Cys Ser Glu Glu Asp Asp 2425 2430 2435	7651
acc acc gtg tgc tgc tcc atg tca tac tcc tgg acc ggg gct cta ata Thr Thr Val Cys Cys Ser Met Ser Tyr Ser Trp Thr Gly Ala Leu Ile 2440 2445 2450	7699
act ccc tgt agc ccc gaa gag gaa aag ttg cca atc aac cct ttg agt Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro Ile Asn Pro Leu Ser 2455 2460 2465	7747
aac tcg ctg ttg cga tac cat aac aag gtg tac tgt aca aca tca aag Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr Cys Thr Thr Ser Lys 2470 2485	7795
agc gcc tca cag agg gct aaa aag gta act ttt gac agg acg caa gtg Ser Ala Ser Gln Arg Ala Lys Lys Val Thr Phe Asp Arg Thr Gln Val 2490 2495 2500	7843
ctc gac gcc cat tat gac tca gtc tta aag gac atc aag cta gcg gct	7891

Leu Asr		His 2505	Tyr	Asp	Ser		Leu 2510	Lys	Asp	Ile		Leu 2515	Ala	Ala	
tcc aac Ser Lys		_	_		Leu			_		Glu	-				7939
act cca Thr Pro 253) Pro			Āla	_		-		Gly						7987
gtc cgo Val Aro 2550	_	_	Ser			_	_	Asn					Val		8035
aag gad Lys As _l		Leu	_	-			Thr					Thr			8083
gcc aaa Ala Lya	s Asn					Val					Gly				8131
cca gc	_			_	Tyr		-			Val		-	_		8179
aaa ate Lys Me ⁻ 261	Āla			Āsp				-	Leu		_		_	-	8227
gga gc Gly Ala 2630			Gly		_			Pro	_			-	Glu		8275
ctc tte Leu Le	-	Āla			_	-	Lys	-		_		Phe	_		8323
gat ace Asp Th	r Arg					Thr					Asp				8371
gag gad Glu Gl	_			_	Āla	_		_		Glu		-			8419
gcc ata Ala Ile 269	e His			Thr					Val						8467
aac ag Asn Se 2710			Gln					Arg					Ser		8515
gtg cta Val Le															8563

2730 2735 2740

cta gcg gcc tgc aag gct gcg ggg ata gtt gcg ccc aca atg ctg gta Leu Ala Ala Cys Lys Ala Ala Gly Ile Val Ala Pro Thr Met Leu Val 2745 2750 2755	8611
tgc ggc gat gac cta gta gtc atc tca gaa agc cag ggg act gag gag Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser Gln Gly Thr Glu Glu 2760 2765 2770	8659
gac gag cgg aac ctg aga gcc ttc acg gag gcc atg acc agg tac tct Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala Met Thr Arg Tyr Ser 2775 2780 2785	8707
gcc cct cct ggt gat ccc ccc aga ccg gaa tat gac ctg gag cta ata Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr Asp Leu Glu Leu Ile 2790 2795 2800 2805	8755
aca tee tgt tee tea aat gtg tet gtg geg ttg gge eeg egg gge ege Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu Gly Pro Arg Gly Arg 2810 2815 2820	8803
cgc aga tac tac ctg acc aga gac cca acc act cca ctc gcc cgg gct Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr Pro Leu Ala Arg Ala 2825 2830 2835	8851
gcc tgg gaa aca gtt aga cac tcc cct atc aat tca tgg ctg gga aac Ala Trp Glu Thr Val Arg His Ser Pro Ile Asn Ser Trp Leu Gly Asn 2840 2845 2850	8899
atc atc cag tat gct cca acc ata tgg gtt cgc atg gtc cta atg aca Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg Met Val Leu Met Thr 2855 2860 2865	8947
cac ttc ttc tcc att ctc atg gtc caa gac acc ctg gac cag aac ctc His Phe Phe Ser Ile Leu Met Val Gln Asp Thr Leu Asp Gln Asn Leu 2870 2875 2880 2885	8995
aac ttt gag atg tat gga tca gta tac tcc gtg aat cct ttg gac ctt Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val Asn Pro Leu Asp Leu 2890 2895 2900	9043
cca gcc ata att gag agg tta cac ggg ctt gac gcc ttt tct atg cac Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp Ala Phe Ser Met His 2905 2910 2915	9091
aca tac tct cac cac gaa ctg acg cgg gtg gct tca gcc ctc aga aaa Thr Tyr Ser His His Glu Leu Thr Arg Val Ala Ser Ala Leu Arg Lys 2920 2925 2930	9139
ctt ggg gcg cca ccc ctc agg gtg tgg aag agt cgg gct cgc gca gtc Leu Gly Ala Pro Pro Leu Arg Val Trp Lys Ser Arg Ala Arg Ala Val 2935 2940 2945	9187
agg gcg tcc ctc atc tcc cgt gga ggg aaa gcg gcc gtt tgc ggc cga Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala Ala Val Cys Gly Arg 2950 2955 2960 2965	9235

tat ctc ttc aat tgg gcg gtg aag acc aag ctc aaa ctc act cca ttg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu Lys Leu Thr Pro Leu 2970 2975 2980	9283
ccg gag gcg cgc cta ctg gac tta tcc agt tgg ttc acc gtc ggc gcc Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp Phe Thr Val Gly Ala 2985 2990 2995	9331
ggc ggg ggc gac att ttt cac agc gtg tcg cgc gcc cga ccc cgc tca Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg Ala Arg Pro Arg Ser 3000 3005 3010	9379
tta ctc ttc ggc cta ctc cta ctt ttc gta ggg gta ggc ctc ttc cta Leu Leu Phe Gly Leu Leu Leu Leu Phe Val Gly Val Gly Leu Phe Leu 3015 3020 3025	9427
ctc ccc gct cgg tag agcggcacac actaggtaca ctccatagct aactgttcct Leu Pro Ala Arg 3030	9482
tttttttt tttttttt tttttttt tttttttt tttt	c 9542
tetttettee etteteatet tattetaett tetttett	g 9602
tcacggctag ctgtgaaagg tccgtgagcc gcatgactgc agagagtgcc gtaactggt	c 9662
tctctgcaga tcatgt	9678
<210> 4 <211> 3033 <212> PRT <213> Hepatitis C virus	
<211> 3033 <212> PRT <213> Hepatitis C virus <400> 4	
<pre><211> 3033 <212> PRT <213> Hepatitis C virus <400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn</pre>	
<211> 3033 <212> PRT <213> Hepatitis C virus <400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn	
<pre><211> 3033 <212> PRT <213> Hepatitis C virus <400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn</pre>	
<pre><211> 3033 <212> PRT <213> Hepatitis C virus </pre> <pre><400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn 1</pre>	
<pre><211> 3033 <212> PRT <213> Hepatitis C virus </pre> <pre><400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn 1</pre>	
<pre><211> 3033 <212> PRT <213> Hepatitis C virus </pre> <pre><400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn 1</pre>	
<pre><211> 3033 <212> PRT <213> Hepatitis C virus </pre> <pre><400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn</pre>	
<pre><211> 3033 <212> PRT <213> Hepatitis C virus </pre> <pre><400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn</pre>	
<pre><211> 3033 <212> PRT <213> Hepatitis C virus </pre> <pre><400> 4 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn</pre>	

```
Gly Val Asn Tyr Ala Thr Gly Asn Leu Pro Gly Phe Pro Phe Ser Ile
                165
                                    170
Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Val Pro Val Ser Ala Ala
           180
                                185
Gln Val Lys Asn Thr Ser Ser Ser Tyr Met Val Thr Asn Asp Cys Ser
                            200
       195
Asn Asp Ser Ile Thr Trp Gln Leu Glu Ala Ala Val Leu His Val Pro
                       215
                                            220
Gly Cys Val Pro Cys Glu Arg Val Gly Asn Thr Ser Arg Cys Trp Val
                                        235
                   230
Pro Val Ser Pro Asn Met Ala Val Arg Gln Pro Gly Ala Leu Thr Gln
                                    250
                245
Gly Leu Arg Thr His Ile Asp Met Val Val Met Ser Ala Thr Phe Cys
           260
                                265
Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ala
                           280
                                                285
Gln Val Phe Ile Val Ser Pro Gln Tyr His Trp Phe Val Gln Glu Cys
                        295
Asn Cys Ser Ile Tyr Pro Gly Thr Ile Thr Gly His Arg Met Ala Trp
                    310
                                        315
Asp Met Met Met Asn Trp Ser Pro Thr Ala Thr Met Ile Leu Ala Tyr
                325
                                    330
Val Met Arq Val Pro Glu Val Ile Ile Asp Ile Val Ser Gly Ala His
            340
                                345
Trp Gly Val Met Phe Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp
                           360
Ala Lys Val Ile Val Ile Leu Leu Leu Ala Ala Gly Val Asp Ala Gly
                        375
                                           380
Thr Thr Thr Val Gly Gly Ala Val Ala Arg Ser Thr Asn Val Ile Ala
                    390
                                        395
Gly Val Phe Ser His Gly Pro Gln Gln Asn Ile Gln Leu Ile Asn Thr
               405
                                    410
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
            420
                                425
Leu Asn Thr Gly Phe Leu Ala Ala Leu Phe Tyr Thr Asn Arg Phe Asn
                            440
Ser Ser Gly Cys Pro Gly Arg Leu Ser Ala Cys Arg Asn Ile Glu Ala
                                            460
                        455
Phe Arg Ile Gly Trp Gly Thr Leu Gln Tyr Glu Asp Asn Val Thr Asn
                                        475
Pro Glu Asp Met Arg Pro Tyr Cys Trp His Tyr Pro Pro Lys Pro Cys
                                    490
                485
Gly Val Val Pro Ala Arg Ser Val Cys Gly Pro Val Tyr Cys Phe Thr
            500
                                505
Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Arg Gly Val Pro Thr
                            520
                                                525
Tyr Thr Trp Gly Glu Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr
                        535
                                            540
Arg Pro Pro Gln Gly Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr
                    550
                                        555
Gly Phe Thr Lys Thr Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp
                                    570
                565
Phe Asn Ala Ser Thr Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys
                                585
His Pro Asp Ala Thr Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr
                        600
Pro Lys Cys Leu Val His Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys
```

```
615
Thr Val Asn Phe Thr Ile Phe Lys Ile Arg Met Tyr Val Gly Gly Val
                   630
                                      635
Glu His Arg Leu Thr Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys
               645
                                   650
Asp Leu Glu Asp Arg Asp Arg Ser Gln Leu Ser Pro Leu Leu His Ser
                              665
Thr Thr Glu Trp Ala Ile Leu Pro Cys Thr Tyr Ser Asp Leu Pro Ala
                          680
Leu Ser Thr Gly Leu Leu His Leu His Gln Asn Ile Val Asp Val Gln
                      695
                                          700
Tyr Met Tyr Gly Leu Ser Pro Ala Ile Thr Lys Tyr Val Val Arg Trp
                  710
                                      715
Glu Trp Val Val Leu Leu Phe Leu Leu Leu Ala Asp Ala Arg Val Cys
               725
                                  730
Ala Cys Leu Trp Met Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu
           740
                              745
Glu Lys Leu Val Val Leu His Ala Ala Ser Ala Ala Asn Cys His Gly
                           760
Leu Leu Tyr Phe Ala Ile Phe Phe Val Ala Ala Trp His Ile Arg Gly
                       775
Arg Val Val Pro Leu Thr Thr Tyr Cys Leu Thr Gly Leu Trp Pro Phe
                  790
                                      795
Cys Leu Leu Met Ala Leu Pro Arg Gln Ala Tyr Ala Tyr Asp Ala
                                  810
               805
Pro Val His Gly Gln Ile Gly Val Gly Leu Leu Ile Leu Ile Thr Leu
                              825
          820
Phe Thr Leu Thr Pro Gly Tyr Lys Thr Leu Leu Gly Gln Cys Leu Trp
                          840
                                              845
Trp Leu Cys Tyr Leu Leu Thr Leu Gly Glu Ala Met Ile Gln Glu Trp
                       855
Val Pro Pro Met Gln Val Arg Gly Gly Arg Asp Gly Ile Ala Trp Ala
                                      875
                   870
Val Thr Ile Phe Cys Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu
               885
                                   890
Leu Ala Leu Leu Gly Pro Ala Tyr Leu Leu Arg Ala Ala Leu Thr His
                               905
Val Pro Tyr Phe Val Arg Ala His Ala Leu Ile Arg Val Cys Ala Leu
                          920
Val Lys Gln Leu Ala Gly Gly Arg Tyr Val Gln Val Ala Leu Leu Ala
                      935
                                         940
Leu Gly Arg Trp Thr Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met
                   950
                                      955
Ser Asp Trp Ala Ala Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu
               965
                                  970
Pro Ile Ile Phe Ser Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala
           980
                               985
Glu Thr Ala Ala Cys Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala
                                1005
                          1000
Arg Leu Gly Gln Glu Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser
                      1015
                                         1020
Lys Gly Trp Lys Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr
        1030
                                  1035
Arg Gly Leu Leu Gly Ala Ile Val Val Ser Met Thr Gly Arg Asp Arg
             1045
                               1050
Thr Glu Gln Ala Gly Glu Val Gln Ile Leu Ser Thr Val Ser Gln Ser
                             1065
```

```
Phe Leu Gly Thr Thr Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly
                   1080
                                  1085
    1075
Ala Gly Asn Lys Thr Leu Ala Gly Leu Arg Gly Pro Val Thr Gln Met
                1095
                               1100
Tyr Ser Ser Ala Glu Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly
1105 1110
                            1115 1120
Thr Lys Ser Leu Glu Pro Cys Lys Cys Gly Ala Val Asp Leu Tyr Leu
          1125 1130 1135
Val Thr Arg Asn Ala Asp Val Ile Pro Ala Arg Arg Gly Asp Lys
   1140 1145 1150
Arg Gly Ala Leu Leu Ser Pro Arg Pro Ile Ser Thr Leu Lys Gly Ser
 1155 1160 1165
Ser Gly Gly Pro Val Leu Cys Pro Arg Gly His Val Val Gly Leu Phe
 1170 1175 1180
Arg Ala Ala Val Cys Ser Arg Gly Val Ala Lys Ser Ile Asp Phe Ile
             1190
                            1195
Pro Val Glu Thr Leu Asp Val Val Thr Arg Ser Pro Thr Phe Ser Asp
                         1210
          1205
Asn Ser Thr Pro Pro Ala Val Pro Gln Thr Tyr Gln Val Gly Tyr Leu
       1220
                      1225
                                     1230
His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro Val Ala Tyr
                   1240
Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser Val Ala Ala
 1250 1255 1260
Thr Leu Gly Phe Gly Ala Tyr Leu Ser Lys Ala His Gly Ile Asn Pro
   1270 1275 1280
Asn Ile Arg Thr Gly Val Arg Thr Val Met Thr Gly Glu Ala Ile Thr
         1285 1290 1295
Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys Ala Ser Gly
  1300 1305
Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ala Val Asp Ala Thr
    1315 1320
                                  1325
Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Ala Glu Thr Ala Gly
                1335
                               1340
Val Arg Leu Thr Val Leu Ala Thr Ala Thr Pro Pro Gly Ser Val Thr
1345 1350
                          1355
Thr Pro His Pro Asp Ile Glu Glu Val Gly Leu Gly Arg Glu Gly Glu
          1365 1370 1375
Ile Pro Phe Tyr Gly Arg Ala Ile Pro Leu Ser Cys Ile Lys Gly Gly
     1380 1385 1390
Arg His Leu Ile Phe Cys His Ser Lys Lys Lys Cys Asp Glu Leu Ala
 1395 1400 1405
Ala Ala Leu Arg Gly Met Gly Leu Asn Ala Val Ala Tyr Tyr Arg Gly
 1410 1415
                               1420
Leu Asp Val Ser Ile Ile Pro Ala Gln Gly Asp Val Val Val Ala
              1430
                            1435
Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp Ser Val Ile
                         1450 1455
          1445
Asp Cys Asn Val Ala Val Thr Gln Ala Val Asp Phe Ser Leu Asp Pro
       1460
                      1465
                                     1470
Thr Phe Thr Ile Thr Thr Gln Thr Val Pro Gln Asp Ala Val Ser Arg
    1475 1480
Ser Gln Arg Arg Gly Arg Thr Gly Arg Gly Arg Gln Gly Thr Tyr Arg
 1490 1495 1500
Tyr Val Ser Thr Gly Glu Arg Ala Ser Gly Met Phe Asp Ser Val Val
1505 1510 1515 1520
Leu Cys Glu Cys Tyr Asp Ala Gly Ala Ala Trp Tyr Asp Leu Thr Pro
```

15	25	1	.530		1535
Ala Glu Thr Thr V	al Arg Leu	Arg Ala 1545	Tyr Phe	Asn Thr	Pro Gly Leu 1550
Pro Val Cys Gln A 1555		Glu Phe 1560	Trp Glu	Ala Val	
Leu Thr His Ile A 1570	sp Ala His 1575			Thr Lys 580	Gln Ala Gly
Glu Asn Phe Ala T 1585	1590	_	1595		1600
Ala Lys Ala Pro P 16	05	1	610		1615
Arg Leu Lys Pro T		1625			1630
Gly Pro Ile Thr A 1635		1640		1645	
Ile Ala Thr Cys M 1650	et Gln Ala 1655			Met Thr .660	Ser Thr Trp
Val Leu Ala Gly G 1665	1670		1675		1680
Thr Gly Cys Val S 16	85	1	.690		1695
Val Val Ala Pro A 1700		1705			1710
Glu Glu Cys Ala S 1715		1720		1725	
Ala Glu Met Leu L 1730	ys Ser Lys 1735			Leu Gln .740	Gln Ala Ser
Lys Gln Ala Gln A 1745	1750		1755		1760
Val Glu Gln Phe T 17	65	1	.770		1775
Gln Tyr Leu Ala G 1780		1785			1790
Ser Met Met Ala P 1795		1800		1805	
Thr Thr Ile Leu L 1810	1815		1	820	
Ala Pro Pro Ala G 1825	1830	_	1835		1840
Ala Ala Val Gly S 18	45	1	.850		1855
Ala Gly Tyr Gly A 1860		1865			1870
Met Ser Gly Glu L 1875		1880	-	1885	
Gly Ile Leu Ser P	1895		1	900	_
Ile Leu Arg Arg H 1905	1910	_	1915		1920
Asn Arg Leu Ile A	25	1	.930		1935
His Tyr Val Thr G	_	1945	_		1950
Gly Ser Leu Thr I 1955		1960		1965	
Thr Glu Asp Cys P 1970	ro Ile Pro 1975			Trp Leu .980	Arg Asp Val

Trp Asp Trp Val Cys Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr 1995 2000 1990 1985 Ser Lys Leu Phe Pro Lys Leu Pro Gly Leu Pro Phe Ile Ser Cys Gln 2010 2015 2005 Lys Gly Tyr Lys Gly Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg 2020 2025 2030 Cys Pro Cys Gly Ala Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met 2035 2040 2045 Arg Ile Thr Gly Pro Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe 2050 2055 2060 Pro Ile Asn Cys Tyr Thr Glu Gly Gln Cys Ala Pro Lys Pro Pro Thr 2065 2070 2075 2080 Asn Tyr Lys Thr Ala Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu 2085 2090 Val Thr Gln His Gly Ser Tyr Ser Tyr Val Thr Gly Leu Thr Thr Asp 2105 2110 Asn Leu Lys Ile Pro Cys Gln Leu Pro Ser Pro Glu Phe Phe Ser Trp 2115 2120 2125 Val Asp Gly Val Gln Ile His Arg Phe Ala Pro Thr Pro Lys Pro Phe 2135 Phe Arg Asp Glu Val Ser Phe Cys Val Gly Leu Asn Ser Tyr Ala Val 2145 2150 2155 2160 Gly Ser Gln Leu Pro Cys Glu Pro Glu Pro Asp Ala Asp Val Leu Arg 2165 2170 2175 Ser Met Leu Thr Asp Pro Pro His Ile Thr Ala Glu Thr Ala Ala Arg 2180 2185 Arg Leu Ala Arg Gly Ser Pro Pro Ser Glu Ala Ser Ser Ser Val Ser 2200 2205 Gln Leu Ser Ala Pro Ser Leu Arg Ala Thr Cys Thr Thr His Ser Asn 2215 2220 Thr Tyr Asp Val Asp Met Val Asp Ala Asn Leu Leu Met Glu Gly Gly 2230 2235 Val Ala Gln Thr Glu Pro Glu Ser Arg Val Pro Val Leu Asp Phe Leu 2245 2250 2255 Glu Pro Met Ala Glu Glu Glu Ser Asp Leu Glu Pro Ser Ile Pro Ser 2260 2265 2270 Glu Cys Met Leu Pro Arg Ser Gly Phe Pro Arg Ala Leu Pro Ala Trp 2275 2280 2285 Ala Arg Pro Asp Tyr Asn Pro Pro Leu Val Glu Ser Trp Arg Arg Pro 2290 2295 2300 Asp Tyr Gln Pro Pro Thr Val Ala Gly Cys Ala Leu Pro Pro Pro Lys 2305 2310 2315 2320 Lys Ala Pro Thr Pro Pro Pro Arg Arg Arg Thr Val Gly Leu Ser 2325 2330 2335 Glu Ser Thr Ile Ser Glu Ala Leu Gln Gln Leu Ala Ile Lys Thr Phe 2345 2340 2350 Gly Gln Pro Pro Ser Ser Gly Asp Ala Gly Ser Ser Thr Gly Ala Gly 2360 2365 Ala Ala Glu Ser Gly Gly Pro Thr Ser Pro Gly Glu Pro Ala Pro Ser 2375 2380 Glu Thr Gly Ser Ala Ser Ser Met Pro Pro Leu Glu Gly Glu Pro Gly 2385 2390 2395 2400 Asp Pro Asp Leu Glu Ser Asp Gln Val Glu Leu Gln Pro Pro Gln 2405 2410 2415 Gly Gly Gly Val Ala Pro Gly Ser Gly Ser Gly Ser Trp Ser Thr Cys 2420 2425

Ser Glu Glu Asp Asp Thr Thr Val Cys Cys Ser Met Ser Tyr Ser Trp 2445 2435 2440 Thr Gly Ala Leu Ile Thr Pro Cys Ser Pro Glu Glu Glu Lys Leu Pro 2455 2460 Ile Asn Pro Leu Ser Asn Ser Leu Leu Arg Tyr His Asn Lys Val Tyr 2465 2470 2475 2480 Cys Thr Thr Ser Lys Ser Ala Ser Gln Arg Ala Lys Lys Val Thr Phe 2485 2490 2495 Asp Arg Thr Gln Val Leu Asp Ala His Tyr Asp Ser Val Leu Lys Asp 2500 2505 2510 Ile Lys Leu Ala Ala Ser Lys Val Ser Ala Arg Leu Leu Thr Leu Glu 2515 2520 2525 Glu Ala Cys Gln Leu Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly 2530 2535 2540 Phe Gly Ala Lys Glu Val Arg Ser Leu Ser Gly Arg Ala Val Asn His 2545 2550 2555 2560 Ile Lys Ser Val Trp Lys Asp Leu Leu Glu Asp Pro Gln Thr Pro Ile 2565 2570 Pro Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala 2580 2585 2590 Lys Gly Gly Lys Lys Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly 2600 2595 2605 Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Ile Thr Gln Lys Leu 2615 2620 Pro Gln Ala Val Met Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala 2625 2630 2635 Gln Arg Val Glu Tyr Leu Leu Lys Ala Trp Ala Glu Lys Lys Asp Pro 2645 2650 2655 Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu 2660 2665 2670 Arg Asp Ile Arg Thr Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro 2680 2685 Glu Glu Ala Arg Thr Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val 2695 2700 Gly Gly Pro Met Phe Asn Ser Lys Gly Gln Thr Cys Gly Tyr Arg Arg 2705 2710 2715 2720 Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr 2725 2730 Cys Tyr Val Lys Ala Leu Ala Ala Cys Lys Ala Ala Gly Ile Val Ala 2740 2745 2750 Pro Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser 2755 2760 2765 Gln Gly Thr Glu Glu Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala 2770 2775 2780 Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr 2790 2795 Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu 2805 2810 Gly Pro Arg Gly Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr 2820 2825 2830 Pro Leu Ala Arg Ala Ala Trp Glu Thr Val Arg His Ser Pro Ile Asn 2840 2845 Ser Trp Leu Gly Asn Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg 2855 2860 Met Val Leu Met Thr His Phe Phe Ser Ile Leu Met Val Gln Asp Thr 2865 2870 2875 2880 Leu Asp Gln Asn Leu Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val

2885 2890 2895	
Asn Pro Leu Asp Leu Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp 2900 2905 2910	
Ala Phe Ser Met His Thr Tyr Ser His His Glu Leu Thr Arg Val Ala 2915 2920 2925	
Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Val Trp Lys Ser 2930 2935 2940	
Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Lys Ala 2945 2950 2955 2960	
Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu 2965 2970 2975	
Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp 2980 2985 2990	
Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Phe His Ser Val Ser Arg 2995 3000 3005	
Ala Arg Pro Arg Ser Leu Leu Phe Gly Leu Leu Leu Phe Val Gly 3010 3015 3020	
Val Gly Leu Phe Leu Leu Pro Ala Arg 3025 3030	
3023	
<210> 5	
<211> 9674 <212> DNA	
<213> Hepatitis C virus	
<220> <221> CDS	
<222> (341)(9442)	
<400> 5 accegecet aataggggeg acaeteegee atgaateact eccetgtgag gaactaetgt	60
cttcacgcag aaagcgtcta gccatggcgt tagtatgagt gtcgtacagc ctccaggccc	
	120
cccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg	
cccctcccg ggagagccat agtggtctgc ggaaccggtg agtacaccgg aattgccggg aagactgggt cctttcttgg ataaacccac tctatgcccg gccatttggg cgtgccccg	180
aagactgggt cetttettgg ataaacceae tetatgeeeg gecatttggg egtgeeeeg	180 240
aagactgggt cetttettgg ataaacccac tetatgeeg gecatttggg egtgeeeeg caagactget ageegagtag egttgggttg egaaaggeet tgtggtactg eetgataggg	180 240 300
aagactggt cettettgg ataaacccac tetatgeeg gecatttggg egtgeeceeg caagactget ageegagtag egttgggttg egaaaggeet tgtggtactg eetgataggg tgettgegag tgeeceggga ggtetegtag acegtgeace atg age aca aat eee Met Ser Thr Asn Pro	180 240
aagactggt cctttcttgg ataaacccac tctatgcccg gccatttggg cgtgcccccg caagactgct agccgagtag cgttgggttg cgaaaggcct tgtggtactg cctgataggg tgcttgcgag tgccccggga ggtctcgtag accgtgcacc atg agc aca aat ccc Met Ser Thr Asn Pro 1 5	180 240 300 355
aagactgggt cetttettgg ataaacccae tetatgeeeg gecatttggg egtgeeeegg caagactget ageegagtag egttgggttg egaaaggeet tgtggtaetg eetgataggg tgettgegag tgeeeeggga ggtetegtag acegtgeace atg age aca aat eee Met Ser Thr Asn Pro 1 5 aaa eet caa aga aaa ace aaa aga aac act aac egt ege eea eaa gae Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp	180 240 300
aagactgggt cetttettgg ataaacceae tetatgeeeg gecatttggg egtgeeeegg caagactget ageegagtag egttgggttg egaaaggeet tgtggtactg eetgataggg tgettgegag tgeeeeggga ggtetegtag acegtgeaee atg age aca aat eee Met Ser Thr Asn Pro 1 5 aaa eet caa aga aaa ace aaa aga aac act aac egt ege eea caa gae	180 240 300 355
aagactgggt cetttettgg ataaacceae tetatgeeeg gecatttggg egtgeeeeeg caagactget ageegagtag egttgggttg egaaaggeet tgtggtactg eetgataggg tgettgegag tgeeeeggga ggtetegtag acegtgeaee atg age aca aat eee Met Ser Thr Asn Pro 1 5 aaa eet caa aga aaa ace aaa aga aac act aac egt ege eea eaa gae Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp 10 15 20 gtt aag ttt eeg gge gge eag ate gtt gge gga gta tac ttg ttg Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu	180 240 300 355
aagactgggt cetttettgg ataaacccac tetatgeeeg gecatttggg egtgeeeeeg caagactget ageegagtag egttgggttg egaaaggeet tgtggtactg eetgataggg tgettgegag tgeeeeggga ggtetegtag acegtgeace atg age aca aat eee Met Ser Thr Asn Pro 1 5 aaa eet caa aga aaa ace aaa aga aac act aac egt ege eea caa gae Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp 10 15 20 gtt aag ttt eeg gge gge gge eag ate gtt gge gga gta tac ttg ttg	180 240 300 355 403
aagactgggt cetttettgg ataaacceae tetatgeeeg gecatttggg egtgeeeeeg caagactget ageegagtag egttgggttg egaaaggeet tgtggtactg eetgataggg tgettgegag tgeeeeggga ggtetegtag acegtgeaee atg age aca aat eee Met Ser Thr Asn Pro 1 5 aaa eet caa aga aaa ace aaa aga aac act aac egt ege eea eaa gae Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn Arg Arg Pro Gln Asp 10 15 20 gtt aag ttt eeg gge gge eag ate gtt gge gga gta tac ttg ttg Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly Gly Val Tyr Leu Leu	180 240 300 355 403

			_		cgt Arg			_	_							547
					tcc Ser 75											595
					ctc Leu			-				_			_	643
					tgg Trp											691
			_	-	atc Ile	-			_	-			-	_		739
_			_		gtc Val	_		_	_		_		-	_	-	787
					gtg Val 155											835
					ggt Gly	-						_	_	_		883
					act Thr		-		-	_			_			931
					gcg Ala			_	_			-	_			979
					gca Ala											1027
		_			aca Thr 235							_				1075
					cct Pro											1123
	-	_	_		ttg Leu		_	_		_		-				1171
ggg	gac	ctc	tgt	ggc	ggg	gtg	atg	ctc	gcg	tcc	cag	atg	ttc	att	gtc	1219

C	Sly	Asp	Leu 280	Cys	Gly	Gly	Val	Met 285	Leu	Ala	Ser	Gln	Met 290	Phe	Ile	Val	
			cag Gln														1267
E			gcc Ala														1315
		-	ccc Pro	_			_		_				_	_	_		1363
			atc Ile														1411
			gcc Ala 360														1459
			ctg Leu														1507
S			gct Ala														1555
			cgg Arg											_			1603
			cgc Arg														1651
			gcc Ala 440														1699
			ctg Leu														1747
C			ctg Leu														1795
			tgc Cys														1843
			gtg Val														1891

505		510	515
		gtg cct act tac Val Pro Thr Tyr	
	_	aac agc acc cga Asn Ser Thr Arg 545	
	-	aac tcc act ggc Asn Ser Thr Gly 560	_
2 22 2		aga gct gac ttc Arg Ala Asp Phe 575	-
	Pro Thr Asp Cys	ttt aga aaa cat Phe Arg Lys His 590	
_		tgg ctc acg cca Trp Leu Thr Pro	
		tac cct tgc aca Tyr Pro Cys Thr 625	
		ggg gga gtt gag Gly Gly Val Glu 640	
		gat cgc tgc aac Asp Arg Cys Asn 655	
		ttg cac tcc acc Leu His Ser Thr 670	
		ttg ccc gct ttg Leu Pro Ala Leu	
		gac gta caa tat Asp Val Gln Tyr 705	
		gtt cga tgg gag Val Arg Trp Glu 720	
		agg gtc tgc gcc Arg Val Cys Ala 735	

						gcc Ala										2611
						gct Ala										2659
						tgg Trp 780										2707
						ggc Gly										2755
-	-		-	-	_	tac Tyr	_		-	-					_	2803
			_	_		gta Val	_								_	2851
		_				agc Ser	_		_			_	_			2899
_		_		_		atg Met 860	_	_			-			_	-	2947
						ggc Gly										2995
-		_			-	ata Ile		_								3043
					-	ggt Gly	_		_	_					-	3091
						aga Arg										3139
				-	_	atg Met 940				-						3187
					_	cac His					_	_		_	-	3235

agc Ser	ggc Gly	ctg Leu	cgg Arg	gac Asp 970	ttg Leu	gcg Ala	gtc Val	gct Ala	gtg Val 975	gag Glu	cct Pro	atc Ile	atc Ile	ttc Phe 980	agt Ser	3283
											gag Glu					3331
	Asp					Leu					cga Arg					3379
Ile					Ala					Ser	aag Lys 1025					3427
	Ala			Thr					Gln		cga Arg			Leu		3475
			Val	_	-	_		Arg	_		aca Thr		Gln			3523
		Gln					Val				ttc Phe	Leu				3571
	Ser		_			Thr	-				gct Ala					3619
Leu					Gly					Met	tac Tyr 1105					3667
	Åsp			Gly					Pro		acc Thr			Leu		3715
			Cys					Leu			gtc Val		Arg			3763
		Ile					Arg				cgg Arg	Gly				3811
	Pro					Thr					tcg Ser					3859
Leu					His					Phe	cgg Arg 1185					3907
tct	cgg	ggt	gtg	gct	aag	tcc	ata	gat	ttc	atc	ccc	gtt	gag	acg	ctc	3955

Ser Arg Gly Val Ala	Lys Ser Ile Asp 195	Phe Ile Pro Val G 1200	lu Thr Leu 1205
gac atc gtc acg cgg Asp Ile Val Thr Arg 1210	Ser Pro Thr Phe		
gct gtg ccc cag acc Ala Val Pro Gln Thr 1225			ro Thr Gly
agt gga aaa agc acc Ser Gly Lys Ser Thr 1240			
aaa gtg ctg gtg ctc Lys Val Leu Val Leu 1255			
gcg tac ttg tcc aag Ala Tyr Leu Ser Lys 1270 1			
gtc aga act gtg acg Val Arg Thr Val Thr 1290	Thr Gly Glu Pro		
aaa ttc ctc gcc gat Lys Phe Leu Ala Asp 1305			sp Ile Ile
ata tgc gat gaa tgc Ile Cys Asp Glu Cys 1320			
ggg aca gtc ctt gac Gly Thr Val Leu Asp 1335			
ctg gcc acg gcc acg Leu Ala Thr Ala Thr 1350 1			
ata gag gag gta gcc Ile Glu Glu Val Ala 1370	Leu Gly Gln Glu		
agg gcg ttt ccc ctg Arg Ala Phe Pro Leu 1385	_	Gly Gly Arg His L	-
tgc cac tca aag aaa Cys His Ser Lys Lys 1400			
atg ggc ttg aac gct	gtg gca tat tac	aga ggg ttg gac g	tc tcc ata 4627

Met Gly 1415	Leu	Asn	Ala		Ala 420	Tyr	Tyr	Arg	-	Leu .425	Asp	Val	Ser	Ile	
ata cca Ile Pro 1430			Gly					Val					Leu		4675
acg ggg Thr Gly		Thr					Ser					Asn			4723
gtc acc Val Thr	Gln	-	-	_		Ser	-	-			Phe				4771
aca cag Thr Gln		_	_		Asp					Ser					4819
cgc acg Arg Thr 1495				Arg					Arg						4867
gag cga Glu Arg 1510			Gly					Val					Cys		4915
gac gca Asp Ala		Āla					Leu					Thr			4963
agg ctc Arg Leu	Arg					Thr					Val				5011
cac ctt His Leu					Ala	_				Leu					5059
gct cat Ala His 1575				Gln					Gly						5107
tta gta	gcc	tat	cag	gcc	aca	gtg	tgc	gcc	agg	gcc	aaa	gcg	ccc	CCC	5155
Leu Val 1590	Ala	Tyr		Ala 1595	Thr	Val	Cys		Arg 1600	Ala	Lys	Ala		Pro 1605	
ccg tcc Pro Ser		Asp					Cys					Lys			5203
ctt gtg Leu Val	Gly					Leu					Ser				5251
gag gtc	acc	ctt	aca	cac	ccc	gtg	aca	aaa	tac	atc	gcc	aca	tgc	atg	5299

Glu Val Thr 1640		s Pro Val Thr 1645	Lys Tyr Ile Ala 1650	Thr Cys Met
			acg tgg gtc ctg Thr Trp Val Leu 1665	
		a Ala Tyr Cys	tta gcg acc ggg Leu Ala Thr Gly 1680	
			cga gct gtc gtc Arg Ala Val Val 1695	
	_		gag atg gag gaa Glu Met Glu Glu	-
	ı Leu Leu Ğl	, , , , ,	cgg ata gcc gag Arg Ile Ala Glu 1730	, ,
-			gcc tct aaa cag Ala Ser Lys Gln 1745	
		n Ala Ser Trp	ccc aag atg gag Pro Lys Met Glu 1760	
-		n Phe Ile Ser	ggc att cag tac Gly Ile Gln Tyr 1775	
			gtg gct tcc atg Val Ala Ser Met	
	Leu Thr Se		act agc acc acc Thr Ser Thr Thr 1810	
			caa att gcg cca Gln Ile Ala Pro 1825	
		l Ser Gly Leu	gtg gga gct gct Val Gly Ala Ala 1840	
			atc ctg gca ggg Ile Leu Ala Gly 1855	
			aag atc atg tct Lys Ile Met Ser	

1865 1870 1875

ccc tcc atg gag gat gtc atc aac ttg ctg cct ggg att ctg tct cca Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro Gly Ile Leu Ser Pro 1880 1885 1890	6019
ggt gct ctg gtg gtg gga gtc atc tgc gcg gcc att ctg cgc cgc cat Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala Ile Leu Arg Arg His 1895 1900 1905	6067
gtg gga ccg ggg gaa ggc gcg gtc caa tgg atg aac agg ctt atc gcc Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met Asn Arg Leu Ile Ala 1910 1915 1920 1925	6115
ttc gct tcc aga gga aac cac gtc gcc cct act cac tac gtg acg gag Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr His Tyr Val Thr Glu 1930 1935 1940	6163
tcg gat gcg tcg cag cgt gtc acc'caa ctg ctt ggc tct ctc act ata Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu Gly Ser Leu Thr Ile 1945 1950 1955	6211
act agt cta ctc agg aga ctt cac aac tgg atc act gag gat tgc ccc Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile Thr Glu Asp Cys Pro 1960 1965 1970	6259
ate cca tgc gcc ggc tcg tgg ctc cgc gat gtg tgg gac tgg gtc tgt Ile Pro Cys Ala Gly Ser Trp Leu Arg Asp Val Trp Asp Trp Val Cys 1975 1980 1985	6307
acc atc cta aca gac ttt aag aac tgg ctg acc tcc aag ctg ttc cca Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr Ser Lys Leu Phe Pro 1990 1995 2000 2005	6355
aag atg cct ggc ctc ccc ttt atc tct tgc caa aag ggg tac aag ggc Lys Met Pro Gly Leu Pro Phe Ile Ser Cys Gln Lys Gly Tyr Lys Gly 2010 2015 2020	6403
gtg tgg gcc ggc act ggc atc atg acc aca cga tgc ccc tgc ggc gcc Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg Cys Pro Cys Gly Ala 2025 2030 2035	6451
aac atc tct ggc aac gtc cgc ttg ggc tct atg aga atc aca gga ccc Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met Arg Ile Thr Gly Pro 2040 2045 2050	6499
aaa acc tgc atg aac acc tgg cag ggg acc ttt cct atc aat tgt tat Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe Pro Ile Asn Cys Tyr 2055 2060 2065	6547
aca gaa ggc cag tgc ttg ccg aaa ccc gcg tta aac ttc aag acc gcc Thr Glu Gly Gln Cys Leu Pro Lys Pro Ala Leu Asn Phe Lys Thr Ala 2070 2075 2080 2085	6595
atc tgg aga gtg gcg gcc tca gag tac gcg gaa gtg acg cag cac gga Ile Trp Arg Val Ala Ala Ser Glu Tyr Ala Glu Val Thr Gln His Gly	6643

2090		2095	2100	
tca tat gcc tat ata Ser Tyr Ala Tyr Ile 2105	Thr Gly Leu			
tgc caa ctc ccc tct Cys Gln Leu Pro Ser 2120		Phe Ser Trp Val		
atc cat agg tcc gcc Ile His Arg Ser Ala 2135				
tcg ttc agc gtt ggg Ser Phe Ser Val Gly 2150			Ser Gln Leu	
tgt gac cct gag ccc Cys Asp Pro Glu Pro 2170				
cca tcc cat atc acg Pro Ser His Ile Thr 2185	Ala Glu Ala			
tca ccc cca tct gag Ser Pro Pro Ser Glu 2200		Ser Ala Ser Gln		
tcg ctg cga gcc acc Ser Leu Arg Ala Thr 2215				
atg gtg gat gcc aac Met Val Asp Ala Asn 2230			Ile Arg Ile	
tct gag tcc aaa gtg Ser Glu Ser Lys Val 2250				
gaa gag ggc gac ctt Glu Glu Gly Asp Leu 2265	Glu Pro Ser			
agg aag agg ttc cca Arg Lys Arg Phe Pro 2280		Pro Ala Trp Ala		
aac cca ccg ctt gtg Asn Pro Pro Leu Val 2295				
act gtt gcg ggc tgt Thr Val Ala Gly Cys 2310			Thr Pro Thr	

		agc gag agc acc ata gga Ser Glu Ser Thr Ile Gly 2340	7363
		ttt ggc cag ccc ccc cca Phe Gly Gln Pro Pro Pro 2355	7411
		gac gcc gcc gac tcc ggc Asp Ala Ala Asp Ser Gly 2370	7459
		tcg gag aca ggt tct acc Ser Glu Thr Gly Ser Thr 2385	7507
	eu Glu Gly Glu Pro	ggg gac cca gac ctg gag Gly Asp Pro Asp Leu Glu 2400 2405	7555
		cag ggg ggg gag gca gct Gln Gly Gly Glu Ala Ala 2420	7603
		tgc tcc gag gag gat gac Cys Ser Glu Glu Asp Asp 2435	7651
	=	tgg acc ggg gct cta ata Trp Thr Gly Ala Leu Ile 2450	7699
		cca att aac tcc ttg agc Pro Ile Asn Ser Leu Ser 2465	7747
	yr His Asn Lys Val	tac tgt act aca tca aag Tyr Cys Thr Thr Ser Lys 2480 2485	7795
		ttt gat agg atg caa gtg Phe Asp Arg Met Gln Val 2500	7843
ctc gac gcc tat tat ga		gac atc aag cta gcg gcc	7891
Leu Asp Ala Tyr Tyr As 2505	sp Ser Val Leu Lys 2510	Asp Ile Lys Leu Ala Ala 2515	
2505 tcc aag gtc agc gca ag	2510 gg ctc ctc acc tta	Asp Ile Lys Leu Ala Ala	7939

gtc co Val Ar 2550			Ser					Asn					Val		8035
aag ga Lys As		ı Leu					Thr					Thr			8083
gcc aa Ala Ly						Val					Gly				8131
cca go Pro Al		g Leu			Tyr					Val					8179
aag at Lys Me 261	et Ālā			Asp					Leu						8227
ggg go Gly Al 2630			Gly					Pro					Glu		8275
ctc tt Leu Le		Āla					Arg					Phe			8323
gat ac Asp Th						Thr					Asp				8371
gag ga Glu Gl	-	r Ile		_	Āla	_				Glu					8419
gcc at Ala II 269	le His	_	_	Thr		_			Val		_				8467
aac aq Asn Se 2710			Gln					Arg					Ser		8515
gtg ct Val Le		r Thr					Thr					Val			8563
cta go Leu Al						Gly					Thr				8611
tgc gg Cys Gl		o Asp			Val					Gln					8659
gac ga	ag cg	g aac	ctg	aga	gcc	ttc	acg	gag	gct	atg	acc	agg	tat	tct	8707

i	_	Glu 2775	Arg	Asn	Leu	Arg	Ala 2780	Phe	Thr	Glu		Met 2785	Thr	Arg	Tyr	Ser	
į		Pro			Asp	ccc Pro 2795				Glu					Leu		8755
			_	Ser		aac Asn			Val	_				Gln		_	8803
	_	_	Tyr		_	acc Thr	_	Āsp					Ile	_		_	8851
	-	Trp	_		_	aga Arg	His			_		Ser		_			8899
	Ile		_		_	cca Pro				_	Arg	_	_	_	_		8947
		Phe			Ile	ctc Leu 2875		_	_	Āsp			_	_	Asn		8995
				Met		gga Gly			Tyr					Leu			9043
		-	Ile		-	agg Arg		His			_	_	Phe		_		9091
		Tyr				gaa Glu	Leu					Ser					9139
	Leu					ctc Leu					Ser						9187
ì		Ala			Ile	tcc Ser 2955				Arg					Gly		9235
				Asn		gcg Ala		_	Thr	_				Thr		_	9283
			Ala			ctg Leu		Leu					Thr				9331
						tat Tyr											9379

3000 3005 3010 tta ctc ctt agc cta ctc cta ctt tct gta ggg gta ggc ctc ttc cta 9427 Leu Leu Ser Leu Leu Leu Ser Val Gly Val Gly Leu Phe Leu 3015 3020 3025 9482 ctc ccc gct cga tag agcggcacac attagctaca ctccatagct aactgttcct Leu Pro Ala Arg 3030 tetteeette teatettatt etaetttett tettggtgge teeatettag eeetggteae 9602 ggctagctgt gaaaggtccg tgagccgcat gactgcagag agtgccgtaa ctggtctctc 9662 tgcagatcat gt 9674 <210> 6 <211> 3033 <212> PRT <213> Hepatitis C virus <400> 6 Met Ser Thr Asn Pro Lys Pro Gln Arg Lys Thr Lys Arg Asn Thr Asn

Arg Arg Pro Gln Asp Val Lys Phe Pro Gly Gly Gly Gln Ile Val Gly 20 25 Gly Val Tyr Leu Leu Pro Arg Arg Gly Pro Arg Leu Gly Val Arg Ala 40 Thr Arg Lys Ala Ser Glu Arg Ser Gln Pro Arg Gly Arg Arg Gln Pro 55 Ile Pro Lys His Arg Arg Ser Thr Gly Lys Ser Trp Gly Lys Pro Gly Tyr Pro Trp Pro Leu Tyr Gly Asn Glu Gly Leu Gly Trp Ala Gly Trp 85 90 Leu Leu Ser Pro Arg Gly Ser Arg Pro Ser Trp Gly Pro Asn Asp Pro 105 Arg His Arg Ser Arg Asn Val Gly Lys Val Ile Asp Thr Leu Thr Cys 120 125 Gly Phe Ala Asp Leu Leu Gly Tyr Val Pro Val Val Gly Ala Pro Leu 135 140 Ser Gly Val Ala Ser Ala Leu Ala His Gly Val Arg Val Leu Glu Asp 150 155 Gly Val Asn Phe Ala Thr Gly Asn Leu Pro Gly Cys Ser Phe Ser Ile 165 170 175 Phe Leu Leu Ala Leu Leu Ser Cys Ile Thr Thr Pro Val Ser Ala Val 180 185 Gln Val Lys Asn Thr Ser Asn Ala Tyr Met Ala Thr Asn Asp Cys Ser 200 195 205 Asn Asp Ser Ile Thr Trp Gln Leu Glu Ala Ala Val Leu His Val Pro 215 Gly Cys Val Pro Cys Glu Lys Met Gly Asn Thr Ser Arg Cys Trp Ile 230 235 Pro Val Ser Pro Asn Val Ala Val Arg Gln Pro Gly Ala Leu Thr Arg 245 250

```
Gly Leu Arg Thr His Ile Asp Met Val Val Leu Ser Ala Thr Leu Cys
            260
                                265
Ser Ala Leu Tyr Val Gly Asp Leu Cys Gly Gly Val Met Leu Ala Ser
                           280
Gln Met Phe Ile Val Ser Pro Gln His His Trp Phe Val Gln Glu Cys
                        295
                                            300
Asn Cys Ser Ile Tyr Pro Gly Ala Ile Thr Gly His Arg Met Ala Trp
                    310
                                        315
Asp Met Met Asn Trp Ser Pro Thr Thr Thr Met Ile Leu Ala Tyr
                                    330
                325
Val Met Arg Val Pro Glu Val Ile Ile Asp Ile Ile Ser Gly Ala His
                                345
Trp Gly Val Met Phe Gly Leu Ala Tyr Phe Ser Met Gln Gly Ala Trp
                            360
Ala Lys Val Val Ile Leu Leu Leu Ala Ser Gly Val Asp Ala Tyr
                       375
                                           380
Thr. Thr Thr Gly Ser Ala Ala Gly Arg Thr Thr Ser Ser Leu Ala
                   390
                                        395
Ser Ala Phe Ser Pro Gly Ala Arg Gln Asn Ile Gln Leu Ile Asn Thr
                405
                                    410
Asn Gly Ser Trp His Ile Asn Arg Thr Ala Leu Asn Cys Asn Asp Ser
                                425
Leu His Thr Gly Phe Phe Thr Ala Leu Phe Tyr Ile His Lys Phe Asn
                            440
                                                445
Ser Ser Gly Cys Pro Glu Arg Leu Ser Ala Cys Arg Asn Ile Glu Asp
                       455
Phe Arg Ile Gly Trp Gly Ala Leu Gln Tyr Asp Asp Asn Val Thr Asn
                   470
                                       475
Pro Glu Asp Met Arg Pro Tyr Cys Trp His Tyr Pro Pro Lys Gln Cys
               485
                                   490
Gly Val Val Pro Ala Gly Thr Val Cys Gly Pro Val Tyr Cys Phe Thr
           500
                                505
Pro Ser Pro Val Val Val Gly Thr Thr Asp Arg Leu Gly Val Pro Thr
                            520
Tyr Thr Trp Gly Glu Asn Glu Thr Asp Val Phe Leu Leu Asn Ser Thr
                        535
                                            540
Arg Pro Pro Ser Gly Ser Trp Phe Gly Cys Thr Trp Met Asn Ser Thr
                    550
                                        555
Gly Phe Thr Lys Thr Cys Gly Ala Pro Pro Cys Arg Thr Arg Ala Asp
                                   570
Phe Asn Thr Ser Thr Asp Leu Leu Cys Pro Thr Asp Cys Phe Arg Lys
            580
                                585
His Pro Glu Ala Thr Tyr Ile Lys Cys Gly Ser Gly Pro Trp Leu Thr
                            600
Pro Lys Cys Leu Val Asp Tyr Pro Tyr Arg Leu Trp His Tyr Pro Cys
                       615
                                            620
Thr Val Asn Tyr Ser Thr Phe Lys Ile Arg Met Tyr Val Gly Gly Val
                    630
                                        635
Glu His Arg Leu Met Ala Ala Cys Asn Phe Thr Arg Gly Asp Arg Cys
                645
                                    650
Asn Leu Glu Asp Arg Asp Arg Ser Gln Gln Thr Pro Leu Leu His Ser
                                665
Thr Thr Glu Trp Ala Ile Leu Pro Cys Ser Phe Ser Asp Leu Pro Ala
                            680
                                               685
Leu Ser Thr Gly Leu Leu His Leu His Gln Asn Ile Val Asp Val Gln
                       695
                                           700
Tyr Met Tyr Gly Leu Ser Pro Ala Leu Thr Gln Tyr Ile Val Arg Trp
```

```
705
                 710
                                   715
Glu Trp Val Val Leu Leu Phe Leu Leu Ala Asp Ala Arg Val Cys
                                730
              725
Ala Cys Leu Trp Met Leu Ile Leu Leu Gly Gln Ala Glu Ala Ala Leu
                            745
          740
Glu Lys Leu Val Val Leu His Ala Ala Ser Ala Ala Ser Cys Asn Gly
                        760
                                           765
Phe Leu Tyr Phe Val Ile Phe Leu Val Ala Ala Trp His Ile Lys Gly
                     775
                                       780
Arg Val Val Pro Leu Ala Ala Tyr Ser Leu Thr Gly Leu Trp Pro Phe
                                   795
     790
Cys Leu Leu Leu Ala Leu Pro Gln Gln Ala Tyr Ala Tyr Asp Ala
             805
                             810
Ser Val His Gly Gln Val Gly Ala Ala Leu Leu Val Leu Ile Thr Leu
          820
                            825
Phe Thr Leu Thr Pro Gly Tyr Lys Thr Leu Leu Ser Gln Ser Leu Trp
                         840
                                           845
Trp Leu Cys Tyr Leu Leu Thr Leu Ala Glu Thr Met Val Gln Glu Trp
                     855
Ala Pro Ser Met Gln Ala Arg Gly Gly Arg Asp Gly Ile Ile Trp Ala
                 870
                                    875
Ala Thr Ile Phe Cys Pro Gly Val Val Phe Asp Ile Thr Lys Trp Leu
                                890
Leu Ala Val Leu Gly Pro Gly Tyr Leu Leu Arg Gly Ala Leu Thr Arg
          900
                            905
Val Pro Tyr Phe Val Arg Ala His Ala Leu Leu Arg Met Cys Thr Met
      915
                        920
                                          925
Val Arg His Leu Ala Gly Gly Arg Tyr Val Gln Met Ala Leu Leu Ala
                     935
Leu Gly Arg Trp Thr Gly Thr Tyr Ile Tyr Asp His Leu Thr Pro Met
                 950
                                    955
Ser Asp Trp Ala Ala Ser Gly Leu Arg Asp Leu Ala Val Ala Val Glu
                                970
Pro Ile Ile Phe Ser Pro Met Glu Lys Lys Val Ile Val Trp Gly Ala
          980
                            985
Glu Thr Ala Ala Cys Gly Asp Ile Leu His Gly Leu Pro Val Ser Ala
                       1000
                                         1005
Arg Leu Gly Arg Glu Ile Leu Leu Gly Pro Ala Asp Gly Tyr Thr Ser
                   1015
                                      1020
Lys Gly Trp Lys Leu Leu Ala Pro Ile Thr Ala Tyr Ala Gln Gln Thr
1025 1030 1035
Arg Gly Leu Leu Gly Ser Ile Val Val Ser Met Thr Gly Arg Asp Lys
             1045
                              1050
Thr Glu Gln Ala Gly Glu Val Gln Val Leu Ser Thr Val Thr Gln Ser
                           1065
         1060
                                              1070
Phe Leu Gly Thr Ser Ile Ser Gly Val Leu Trp Thr Val Tyr His Gly
                       1080
                                          1085
Ala Gly Asn Lys Thr Leu Ala Gly Ser Arg Gly Pro Val Thr Gln Met
                    1095
                                      1100
  1090
Tyr Ser Ser Ala Glu Gly Asp Leu Val Gly Trp Pro Ser Pro Pro Gly
                1110
                                  1115
Thr Lys Ser Leu Glu Pro Cys Thr Cys Gly Ala Val Asp Leu Tyr Leu
                              1130 1135
           1125
Val Thr Arg Asn Ala Asp Val Ile Pro Ala Arg Arg Gly Asp Lys
          1140 1145
                                             1150
Arg Gly Ala Leu Leu Ser Pro Arg Pro Leu Ser Thr Leu Lys Gly Ser
```

1155		1160		1165	
Ser Gly Gly Pro	Val Leu Cys 1175			Ala Val	Gly Ile Phe
Arg Ala Ala Val 1185	Cys Ser Arg 1190	Gly Val	Ala Lys 1195	Ser Ile	Asp Phe Ile 1200
Pro Val Glu Thr	Leu Asp Ile 1205		Arg Ser 1210	Pro Thr	Phe Ser Asp 1215
Asn Ser Thr Pro		Pro Gln 1225	Thr Tyr		Gly Tyr Leu 1230
His Ala Pro Thr 1235		Lys Ser 1240	Thr Lys	Val Pro 1245	Val Ala Tyr
Ala Ala Gln Gly 1250	7 Tyr Lys Val 1255			Pro Ser 1260	Val Ala Ala
Thr Leu Gly Phe	1270		1275		1280
Asn Ile Arg Thr	Gly Val Arg 1285		Thr Thr 1290	Gly Glu	Pro Ile Thr 1295
Tyr Ser Thr Tyr 1300		Leu Ala 1305	Asp Gly		Ala Gly Gly 1310
Ala Tyr Asp Ile 1315	_	1320	_	1325	
Thr Ile Leu Gly 1330	v Ile Gly Thr 1335		=	Ala Glu 1340	Thr Ala Gly
Val Arg Leu Thr 1345	Val Leu Ala 1350	Thr Ala	Thr Pro 1355	Pro Gly	Ser Val Thr 1360
Thr Pro His Pro	1365		1370	_	1375
Ile Pro Phe Tyr 1380)	1385			1390
Arg His Leu Ile 1395		1400		1405	
Thr Ala Leu Aro 1410	1415	•	-	1420	
Leu Asp Val Ser 1425	1430		1435		1440
Thr Asp Ala Leu	1445	-	1450		1455
Asp Cys Asn Val)	1465			1470
Thr Phe Thr Ile		1480		1485	
Ser Gln Arg Arg	1495			L500	
Tyr Val Ser Thi	1510		1515		1520
Leu Cys Glu Cys	1525		1530		1535
Val Glu Thr Thi)	1545			1550
Pro Val Cys Glr 1555		1560		1565	
Leu Thr His Ile	1575	•	:	L580 _	
Glu Asn Phe Ala	1590	_	1595		1600
Ala Lys Ala Pro	Pro Pro Ser 1605	_	Val Met 1610	Trp Lys	Cys Leu Thr 1615

```
Arg Leu Lys Pro Thr Leu Val Gly Pro Thr Pro Leu Leu Tyr Arg Leu
              1625
                           1630
Gly Ser Val Thr Asn Glu Val Thr Leu Thr His Pro Val Thr Lys Tyr
    1635 1640 1645
Ile Ala Thr Cys Met Gln Ala Asp Leu Glu Val Met Thr Ser Thr Trp
                1655 1660
Val Leu Ala Gly Gly Val Leu Ala Ala Val Ala Ala Tyr Cys Leu Ala
             1670 1675 1680
Thr Gly Cys Val Ser Ile Ile Gly Arg Leu His Ile Asn Gln Arg Ala
          1685 1690 1695
Val Val Ala Pro Asp Lys Glu Val Leu Tyr Glu Ala Phe Asp Glu Met
      1700 1705 1710
Glu Glu Cys Ala Ser Arg Ala Ala Leu Leu Glu Glu Gly Gln Arg Ile
    1715 1720
                                  1725
Ala Glu Met Leu Lys Ser Lys Ile Gln Gly Leu Leu Gln Gln Ala Ser
 1730 1735 1740
Lys Gln Ala Gln Asp Ile Gln Pro Ala Val Gln Ala Ser Trp Pro Lys
             1750
                            1755
Met Glu Gln Phe Trp Ala Lys His Met Trp Asn Phe Ile Ser Gly Ile
          1765 1770 1775
Gln Tyr Leu Ala Gly Leu Ser Thr Leu Pro Gly Asn Pro Ala Val Ala
                      1785
        1780
Ser Met Met Ala Phe Ser Ala Ala Leu Thr Ser Pro Leu Ser Thr Ser
    1795 1800 1805
Thr Thr Ile Leu Leu Asn Ile Leu Gly Gly Trp Leu Ala Ser Gln Ile
 1810 1815 1820
Ala Pro Pro Ala Gly Ala Thr Gly Phe Val Val Ser Gly Leu Val Gly
1825 1830 1835 1840
Ala Ala Val Gly Ser Ile Gly Leu Gly Lys Val Leu Val Asp Ile Leu
          1845 1850 1855
Ala Gly Tyr Gly Ala Gly Ile Ser Gly Ala Leu Val Ala Phe Lys Ile
       1860 1865 1870
Met Ser Gly Glu Lys Pro Ser Met Glu Asp Val Ile Asn Leu Leu Pro
    1875 · 1880 1885
Gly Ile Leu Ser Pro Gly Ala Leu Val Val Gly Val Ile Cys Ala Ala
                1895
                                1900
Ile Leu Arg Arg His Val Gly Pro Gly Glu Gly Ala Val Gln Trp Met
1905 1910 1915 1920
Asn Arg Leu Ile Ala Phe Ala Ser Arg Gly Asn His Val Ala Pro Thr
         1925 1930
His Tyr Val Thr Glu Ser Asp Ala Ser Gln Arg Val Thr Gln Leu Leu
 1940 1945
                                    1950
Gly Ser Leu Thr Ile Thr Ser Leu Leu Arg Arg Leu His Asn Trp Ile
                                  1965
    1955 1960
Thr Glu Asp Cys Pro Ile Pro Cys Ala Gly Ser Trp Leu Arg Asp Val
                1975
                                1980
Trp Asp Trp Val Cys Thr Ile Leu Thr Asp Phe Lys Asn Trp Leu Thr
              1990
                             1995
Ser Lys Leu Phe Pro Lys Met Pro Gly Leu Pro Phe Ile Ser Cys Gln
           2005
                          2010
                                         2015
Lys Gly Tyr Lys Gly Val Trp Ala Gly Thr Gly Ile Met Thr Thr Arg
                      2025
       2020
                                     2030
Cys Pro Cys Gly Ala Asn Ile Ser Gly Asn Val Arg Leu Gly Ser Met
                   2040
Arg Ile Thr Gly Pro Lys Thr Cys Met Asn Thr Trp Gln Gly Thr Phe
  2050 2055 2060
Pro Ile Asn Cys Tyr Thr Glu Gly Gln Cys Leu Pro Lys Pro Ala Leu
```

2065 2070	207	5 2080
Asn Phe Lys Thr Ala Ile Ti		
2085	2090	2095
Val Thr Gln His Gly Ser Ty 2100	yr Ala Tyr Ile Th 2105	er Gly Leu Thr Thr Asp 2110
Asn Leu Lys Val Pro Cys Gi 2115	ln Leu Pro Ser Pr 2120	o Glu Phe Phe Ser Trp 2125
Val Asp Gly Val Gln Ile H	is Arg Ser Ala Pr	
Phe Arg Asp Glu Val Ser Ph 2145 2150	he Ser Val Gly Le 215	
Gly Ser Gln Leu Pro Cys As		
Ser Met Leu Thr Asp Pro Se	er His Ile Thr Al	a Glu Ala Ala Ala Arg
2180 Arg Leu Ala Arg Gly Ser Pi	2185	2190 a Ser Ser Ser Ala Ser
2195	2200	2205
Gln Leu Ser Ala Pro Ser Le 2210 221		s Thr Thr His Gly Arg 2220
Thr Tyr Asp Val Asp Met Va 2225 2230	al Asp Ala Asn Le 223	
Val Ile Arg Ile Glu Ser Gl 2245		
Asp Ser Met Thr Glu		
Glu Tyr Met Leu Pro Arg Ly		
Ala Arg Pro Asp Tyr Asn Pr 2290 229	ro Pro Leu Val Gl	
Asp Tyr Gln Pro Pro Thr Va	al Ala Gly Cys Al	a Leu Pro Pro Pro Lys
2305 2310 Lys Thr Pro Thr Pro Pro Pro	231 ro Arg Arg Arg Ar	
2325	2330	2335
Glu Ser Thr Ile Gly Asp Ai 2340	2345	2350
Gly Gln Pro Pro Pro Ser Gl 2355	ly Asp Ser Gly Le 2360	u Ser Thr Gly Ala Asp 2365
Ala Ala Asp Ser Gly Asp Asp 2370 237	=	p Glu Leu Ala Leu Ser 2380
Glu Thr Gly Ser Thr Ser Se 2385 2390		eu Glu Gly Glu Pro Gly
Asp Pro Asp Leu Glu Pro G	lu Gln Val Glu Le	u Gln Pro Pro Pro Gln
2405 Gly Gly Glu Ala Ala Pro G	2410 ly Ser Asp Ser Gl	v Ser Tro Ser Thr Cvs
2420	2425	2430
Ser Glu Glu Asp Asp Ser Va 2435	2440	2445
Thr Gly Ala Leu Ile Thr Pr 2450 245		u Glu Glu Lys Leu Pro 2460
Ile Asn Ser Leu Ser Asn Se 2465 2470	er Leu Leu Arg Ty 247	
Cys Thr Thr Ser Lys Ser A	la Ser Leu Arg Al 2490	a Lys Lys Val Thr Phe 2495
Asp Arg Met Gln Val Leu As 2500		
Ile Lys Leu Ala Ala Ser Ly 2515		
2010		2020

```
Glu Ala Cys Gln Leu Thr Pro Pro His Ser Ala Arg Ser Lys Tyr Gly
                2535
Phe Gly Ala Lys Glu Val Arg Ser Leu Ser Gly Arg Ala Val Asn His
             2550
                             2555
Ile Lys Ser Val Trp Lys Asp Leu Leu Glu Asp Ser Gln Thr Pro Ile
           2565
                          2570
Pro Thr Thr Ile Met Ala Lys Asn Glu Val Phe Cys Val Asp Pro Ala
       2580 2585 2590
Lys Gly Gly Lys Lys Pro Ala Arg Leu Ile Val Tyr Pro Asp Leu Gly
    2595 2600 2605
Val Arg Val Cys Glu Lys Met Ala Leu Tyr Asp Val Thr Gln Lys Leu
 2610 2615 2620
Pro Gln Ala Val Met Gly Ala Ser Tyr Gly Phe Gln Tyr Ser Pro Ala
2625 2630 2635 2640
Gln Arg Val Glu Phe Leu Leu Lys Ala Trp Ala Glu Lys Arg Asp Pro
          2645 2650 2655
Met Gly Phe Ser Tyr Asp Thr Arg Cys Phe Asp Ser Thr Val Thr Glu
            2665 2670
        2660
Arg Asp Ile Arg Thr Glu Glu Ser Ile Tyr Gln Ala Cys Ser Leu Pro
                    2680
                                   2685
Glu Glu Ala Arg Thr Ala Ile His Ser Leu Thr Glu Arg Leu Tyr Val
                2695 2700
Gly Gly Pro Met Phe Asn Ser Lys Gly Gln Ser Cys Gly Tyr Arg Arg
2705 2710 2715
Cys Arg Ala Ser Gly Val Leu Thr Thr Ser Met Gly Asn Thr Ile Thr
      2725 2730 2735
Cys Tyr Val Lys Ala Leu Ala Ala Cys Lys Ala Ala Gly Ile Ile Ala
   2740 2745 2750
Pro Thr Met Leu Val Cys Gly Asp Asp Leu Val Val Ile Ser Glu Ser
 2755 2760 2765
Gln Gly Thr Glu Glu Asp Glu Arg Asn Leu Arg Ala Phe Thr Glu Ala
                2775 2780
Met Thr Arg Tyr Ser Ala Pro Pro Gly Asp Pro Pro Arg Pro Glu Tyr
2785 · 2790 *2795 2800
Asp Leu Glu Leu Ile Thr Ser Cys Ser Ser Asn Val Ser Val Ala Leu
          2805 2810
Gly Pro Gln Gly Arg Arg Tyr Tyr Leu Thr Arg Asp Pro Thr Thr
       2820 2825 2830
Ser Ile Ala Arg Ala Ala Trp Glu Thr Val Arg His Ser Pro Val Asn
    2835 2840 2845
Ser Trp Leu Gly Asn Ile Ile Gln Tyr Ala Pro Thr Ile Trp Val Arg
 2850 2855 2860
Met Val Leu Met Thr His Phe Phe Ser Ile Leu Met Ala Gln Asp Thr
                           2875 2880
2865 2870
Leu Asp Gln Asn Leu Asn Phe Glu Met Tyr Gly Ser Val Tyr Ser Val
                          2890
          2885
Ser Pro Leu Asp Leu Pro Ala Ile Ile Glu Arg Leu His Gly Leu Asp
        2900
                       2905
                                      2910
Ala Phe Ser Leu His Thr Tyr Thr Pro His Glu Leu Thr Arg Val Ala
                    2920
                                   2925
Ser Ala Leu Arg Lys Leu Gly Ala Pro Pro Leu Arg Ala Trp Lys Ser
  2930 2935
                                2940
Arg Ala Arg Ala Val Arg Ala Ser Leu Ile Ser Arg Gly Gly Arg Ala
             2950
                             2955
Ala Val Cys Gly Arg Tyr Leu Phe Asn Trp Ala Val Lys Thr Lys Leu
           2965
                         2970
```

Lys Leu Thr Pro Leu Pro Glu Ala Arg Leu Leu Asp Leu Ser Ser Trp 2980 2985 2990 Phe Thr Val Gly Ala Gly Gly Gly Asp Ile Tyr His Ser Val Ser Arg 3005 3000 Ala Arg Pro Arg Leu Leu Leu Ser Leu Leu Leu Ser Val Gly 3015 3020 Val Gly Leu Phe Leu Leu Pro Ala Arg <210> 7 <211> 8024 <212> RNA <213> Artificial Sequence <220> <223> Description of Artificial Sequence: replicon accugecceu aauaggggeg acacueegee augaaucaeu eeecugugag gaacuacugu 60 cuucacgcag aaagcgccua gccauggcgu uaquaugagu gucquacagc cuccaggccc 120 ccccucccq qqaqaqccau aquqqucuqc qqaaccqquq aquacaccqq aauuqccqqq 180 aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240 caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300 cgcuugcgag uqccccqqqa qqucucquaq accquqcacc augaqcacaa auccuaaacc 360

ucaaagaaaa accaaaagaa acaccaaccg ucgcccaaug auugaacaag auggauugca 420 cgcagguucu ccggccqcuu qqquqqaqaq qcuauucqqc uauqacuqqq cacaacaqac 480 aaucqqcuqc ucuqauqccq ccququuccq qcuqucaqcq caqqqqcqcc cqquucuuuu 540 ugucaaqacc qaccuquccg quqcccuqaa uqaacuqcaq qacqaqqcaq cqcqqcuauc 600 quqqcuqqcc acqacqqqcq uuccuuqcqc aqcuquqcuc qacquuquca cuqaaqcqqq 660 aagggacugg cugcuauugg gcgaagugcc ggggcaggau cuccugucau cucaccuugc 720 uccugeegag aaaguaueea ucauggeuga ugeaaugegg eggeugeaua egeuugauee 780 ggcuaccugc ccauucgacc accaagcgaa acaucgcauc gagcgagcac guacucggau 840 ggaagccggu cuugucgauc aggaugaucu ggacgaagag caucaggggc ucgcgccaqc 900 cqaacuquuc qccaqqcuca aqqcqcqcau qccqacqqc qaqqaucucq ucquqaccca 960 uggcgaugcc ugcuugccga auaucauggu ggaaaauggc cgcuuuucug gauucaucga 1020 cuguggccgg cugggugugg cggaccgcua ucaggacaua gcguuggcua cccgugauau 1080 ugcugaagag cuuggcggcg aaugggcuga ccgcuuccuc gugcuuuacg guaucgccgc 1140 ucccgauucg cagcgcaucg ccuucuaucg ccuucuugac gaguucuucu gaguuuaaac 1200 ccucucceuc ccccccccu aacguuacug gccgaagccg cuuggaauaa ggccggugug 1260 cguuugucua uauguuauuu uccaccauau ugccgucuuu uggcaaugug agggcccgga 1320 aaccuqqccc uqucuucuuq acqaqcauuc cuaqqqqucu uuccccucuc qccaaaqqaa 1380 ugcaaqqucu quuqaauquc quqaaqqaaq caquuccucu qqaaqcuucu uqaaqacaaa 1440 caacgucugu agcgacccuu ugcaggcagc ggaaccccc accuggcgac aggugccucu 1500 gcqqccaaaa gccacgugua uaaqauacac cugcaaaggc ggcacaaccc cagugccacg 1560 uuquqaquuq gauaquuquq gaaaqaquca aauggcucuc cucaagcgua uucaacaagg 1620 ggcugaagga ugcccagaag guaccccauu guaugggauc ugaucugggg ccucggugca 1680 caugeuuuac auguquuuag ucgaqquuaa aaaaacqueu aggeeeeeg aaceaegggg 1740 acgugguuuu ccuuugaaaa acacgaugau accauggcuc ccaucacugc uuaugcccag 1800 caaacacgag gccuccuggg cgccauagug gugaguauga cggggcguga caggacagaa 1860 caggeegggg aagueeaaau ceugueeaca gueueueagu eeuueeuegg aacaaceaue 1920 ucggggguuu uguggacugu uuaccacgga gcuggcaaca agacucuagc cggcuuacgg 1980 gguccgguca cgcagaugua cucgagugcu gagggggacu ugguaggcug gcccagcccc 2040 ccugggacca agucuuugga gccgugcaag uguggagccg ucgaccuaua ucuggucacg 2100 cggaacgcug augucauccc ggcucggaga cgcggggaca agcggggagc auugcucucc 2160 cegagaceca uuucgaceuu gaagggguee ueggggggge eggugeueug eecuagggge 2220

						0000
	ggcucuuccg					
uucauccccg	uugagacacu	cgacguuguu	acaaggucuc	ccacuuucag	ugacaacagc	2340
acgccaccgg	cugugcccca	gaccuaucag	gucggguacu	ugcaugcucc	aacuggcagu	2400
ggaaagagca	ccaagguccc	ugucgcguau	gccgcccagg	gguacaaagu	acuagugcuu	2460
	uagcugccac					
	uuaggacugg					
	aauuucucgc					
	gccacgcugu					
	cagccggggu					
	cccaucccga					
	gggcgauucc					
	aaaaguguga					
	auagaggguu					
	acgcccucau					
	ucacccaagc					3120
	cacaagacgc					3180
	cuuauaggua					3240
	gugagugcua					3300
	ggcuuagagc					
cuugaauuuu	gggaggcagu	uuucaccggc	cucacacaca	uagacgccca	cuuccucucc	3420
	aagcggggga					3480
gccagagcca	aggccccucc	cccguccugg	gacgccaugu	ggaagugccu	ggcccgacuc	3540
	uugcgggccc					3600
	cacacccugg					3660
	gcacgugggu					3720
	gaugeguuue					
	aggagguccu					3840
	ucgaagaggg					3900
	aggccucuaa					3960
						4020
	aacaauuuug					4080
	ugucaacacu					4140
	ccaguccguu					4200
	cccagaucgc					
	ccgugggcag					4260
	gcauuucggg					4320
	augucaucaa					4380
	gcgcggccau					4440
	ggcuuauugc					
	cggaugcguc					
	gaagacucca					
	gcgacgugug					
cugaccucua	aauuguuccc	caagcugccc	ggccuccccu	ucaucucuug	ucaaaagggg	4740
uacaagggug	ugugggccgg	cacuggcauc	augaccacgc	gcugcccuug	cggcgccaac	4800
aucucuggca	auguccgccu	gggcucuaug	aggaucacag	ggccuaaaac	cugcaugaac	4860
accuggcagg	ggaccuuucc	uaucaauugc	uacacggagg	gccagugcgc	gccgaaaccc	4920
cccacgaacu	acaagaccgc	caucuggagg	guggcggccu	cggaguacgc	ggaggugacg	4980
cagcaugggu	cguacuccua	uquaacaqqa	cugaccacug	acaaucugaa	aauuccuugc	5040
	cuccagaguu					
	agccguuuuu					
acnancaaan	cccagcuucc	cnanasaccii	gagggggagg	cadacdnami	gaggiiccaiig	5220
cuaacadanc	cgcccacau	cacaacaaaa	achacaacac	aacaciiiiaac	acadadanca	5280
Collegandra	aggcgagcuc	cheadhasac	Cadchancad	Caccanagge	acadaccaca	5340
ucaccacca	acagcaacac	cualidacara	dacallegues	augecaaceu	acucauaaa	5400
adeaccaecc	Cucacacacac	accuarance	aggaigeee	uugeeaaceu	yeucauggag	5460
anage cases	cucagacaga	goodyagacc	ugggugeeeg	aggaggaggu	acucyaycca	5520
auggeegagg	aagagagcga	accaracina	ucaauaccau	cyyagugcau	geneeddagg	5500
aycygguuuc	cacgggccuu	accggcuugg	geacygeeug	acuacaaccc	geegeuegug	2200
	ggaggccaga	111120000000	000000000000	A	110110000000	5610

```
cccaagaagg ccccgacgcc ucccccaagg agacgccgga cagugggucu gagcgagagc 5700
accauaucag aagcccucca gcaacuggec aucaagaccu uuggccagec ecccucgage 5760
qquqauqcaq qcucquccac qqqqqcqqqc qccqccqaau ccggcggucc gacguccccu 5820
ggugagccgg ccccucaga gacagguucc gccuccucua ugcccccccu cgagggggag 5880
ccuggagauc cggaccugga gucugaucag guagagcuuc aaccucccc ccaggggggg 5940
gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000
accgugugeu geuceaugue auaeuceugg accggggeue uaauaaeuce euguageeee 6060
gaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120
guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180
acgcaagugc ucgacgccca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240
aaggucagcg caaggcuccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300
qcaaqaucca aquauqqauu cqqqqccaaq qaqquccqca qcuuquccqq qaqqqccquu 6360
aaccacauca aquecququq qaaqqaccuc cuqqaaqacc cacaaacacc aauucccaca 6420
accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480
gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540
gacauuacac aaaaqcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600
ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660
uuuucquaug auacccgaug cuucqacuca accqucacug agagagacau caggaccgag 6720
gaquecauau accaggecug cueecugeee gaggaggeee geacugeeau acacuegeug 6780
acuqaqaqac uuuacquaqq aqqqcccauq uucaacaqca aqqqucaaac cuqcqquuac 6840
agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900
gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau gcugguaugc 6960
ggcaaugacc uaguagucau cucagaaagc caggggacug aggaggacga gcggaaccug 7020
agagecuuca eggaggecau gaccagguac ucugececue euggugauce ecceagaceg 7080
gaauaugacc uggagcuaau aacauccugu uccucaaaug ugucuguggc guugggcccg 7140
eggggeegee geagauaeua eeugaeeaga gaeeeaaeea euceaeuege eegggeugee 7200
ugggaaacag uuagacacuc cccuaucaau ucauggcugg gaaacaucau ccaguaugcu 7260
ccaaccauau ggguucgcau gguccuaaug acacacuucu ucuccauucu caugguccaa 7320
gacacccugg accagaaccu caacuuugag auguauggau caguauacuc cgugaauccu 7380
uuqqaccuuc caqccauaau uqaqaqquua cacqqqcuuq acqccuuuuc uauqcacaca 7440
uacucucace acgaacugae geggguggeu ucageccuca gaaaacuugg ggegecacee 7500
cucagggugu ggaagagucg ggcucgcgca gucagggcgu cccucaucuc ccguggaggg 7560
aaageggeeg uuugeggeeg auaucucuuc aauugggegg ugaagaccaa geucaaacuc 7620
acuccauuqe eqqaqqeq ccuacuqqac uuauccaquu qquucacequ eqqeqeeqqe 7680
gggggcgaca uuuuucacag cgugucgcgc gcccgacccc gcucauuacu cuucggccua 7740
cuccuacuuu ucquaqqqqu aqqccucuuc cuacuccccq cucqquaqaq cqqcacacac 7800
uuuuuuuuu cuuuuuuuu uuuucccuc uuucuucccu ucucaucuua uucuacuuuc 7920
uuucuuggug gcuccaucuu agcccuaguc acggcuagcu gugaaagguc cgugagccgc 7980
                                                                 8024
augacugcag agagugccgu aacuggucuc ucugcagauc augu
<210> 8
<211> 7994
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: replicon
<400> 8
accuqeeccu aauaqqqqeq acacueeqee auqaaucacu eeccuquqaq qaacuacuqu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccucccq qqaqaqccau aquqqucuqc qqaaccqquq aquacaccqq aauuqccqqq 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacug ccugauaggg 300
cgcuugcgag ugccccggga ggucucguag accgugcacc augagcacaa auccuaaacc 360
```

ucaaadaaaa	accaaaagaa	acaccaacca	ncacccaana	auugaacaag	auggauugca	420
					cacaacagac	
					cgguucuuuu	
			ugaacugcag			600
						660
			agcugugcuc			720
			ggggcaggau			780
			ugcaaugcgg			
					guacucggau	900
			ggacgaagag			
					ucgugaccca	1000
			ggaaaauggc		,	1020
			ucaggacaua		, ,	1080
			ccgcuuccuc		5 5 - 5	1140
			ccuucuugac			1200
			gccgaagccg			1260
			ugccgucuuu		222	1320
			cuaggggucu			1380
			caguuccucu			1440
caacgucugu	agcgacccuu	ugcaggcagc	ggaacccccc	accuggcgac	aggugccucu	1500
gcggccaaaa	gccacgugua	uaagauacac	cugcaaaggc	ggcacaaccc	cagugccacg	1560
uugugaguug	gauaguugug	gaaagaguca	aauggcucuc	cucaagcgua	uucaacaagg	1620
ggcugaagga	ugcccagaag	guaccccauu	guaugggauc	ugaucugggg	ccucggugca	1680
caugcuuuac	auguguuuag	ucgagguuaa	aaaaacgucu	aggccccccg	aaccacgggg	1740
acgugguuuu	ccuuugaaaa	acacgaugau	accauggcuc	ccaucacugc	uuaugcccag	1800
caaacacgag	gccuccuggg	cgccauagug	gugaguauga	cggggcguga	caggacagaa	1860
			gucucucagu			1920
			gcuggcaaca			1980
					gcccagcccc	2040
ccuqqqacca	agucuuugga	gccgugcaag	uquqqaqccq	ucgaccuaua	ucuggucacg	2100
cqqaacqcuq	augucauccc	ggcucggaga	cacaaaaaca	agcggggagc	auugcucucc	2160
ccqaqaccca	uuucqaccuu	gaaggggucc	ncadadadac	cadadcacad	cccuaggggc	2220
cacqucquuq	ggcucuuccg	agcagcugug	uqcucucqqq	gcquqqccaa	auccaucgau	2280
uucaucccca	uuqaqacacu	cgacguuguu	acaaggucuc	ccacuuucag	ugacaacagc	2340
acaccaccaa	cuquqcccca	gaccuaucag	queqqquaeu	ugcaugcucc	aacuggcagu	2400
ggaaagagca	ccaagguccc	ugucacauau	accacccaaa	gguacaaagu	acuagugcuu	2460
aaccccucaa	uagcugccac	ccnaaaannn	adadcanacc	uauccaaggc	acauggcauc	2520
			gugaugaccg			2580
			ugcgcuagcg			2640
					gguccuugau	
					ccccggguca	
didacaaccc	cccauccca	uauagagag	gugeaggeaa	adcadagada	ugagaucccc	2820
niiciialiaaaa	aaacaanncc	ccualiccuae	aucaaggeedeg	ggegggaggg	gauuuucugc	2880
cacucaaaga	agaggaaacc	caaaccage	accaagggag	ggagacaccu	cuugaaugcc	2940
					uguggugguc	3060
			acuggagacu		2 2	3120
			agccuggacc			3180
			agucagcgcc			
			ggugaacgag		, ,	3240
			gcugcguggu			3300
accaccguca	ggcuuagagc	guauuucaac	acgcccggcc	uacccgugug	ucaagaccau	3360
cuugaauuuu	gggaggcagu	uuucaccggc	cucacacaca	uagacgccca	cuuccucucc	3420
					uacggugugc	3480
			gacgccaugu		22	3540
			cuguaccguu			3600
			aucgccacau			3660
			ggaguccugg			3720
cuggcgacug	gaugcguuuc	caucaucggc	cgcuugcacg	ucaaccagcg	agucgucguu	3780

geggecucue ucquaqagqq gaqqqqaaqaq qaaqqqaaqaq qaaqqaaqaqaqaq							
uygʻcugʻcagʻ agʻçcucula gaʻagʻqcagaa agʻquucaqigʻ 3960 cccaaagqga aacaauuuug ggcagaaca gccagaaca augugaacu ucauuagcgg caauuagcgg 4080 gccgcccucia cccaqueggu gccagaaca accacaacu 4080 ugquaqqqu cccaquegga accacaccag ggagcaacag quuquaqqa 4080 uuqqaaqqq cccaquegg accacaccag ggagacacag quuqqaaa cuuqqaaqqa 4260 uuqqaaqaa qqaaqaaca qqaaqaaca 4270 4280 4280 uuqqaaqaa qqaaqaacaa qqaaqaacaa 4290 4280 4280 uuqqaaqaa qqaaqaacaa qqaaqaacaa 4290 4280 4280 uuqqaacaaa qqaaqaagaa qaaqaacaa 4300 480	gcgccggaua	aggagguccu	guaugaggcu	uuugaugaga	uggaggaaug	cgccucuagg	
ccéasaguga aicaeauuuug geefagaaca augugaacau ucucaaggaau duucaacaau decegoguug cuucaaggaa guucaacaau duucaacaau geegagaaca duucaacaau guucaaggag 4100 ugquuaqegu cecagagaaca accaecaaca cuucuaagaa gagagaceu 4200 uauguuqeg gecauuucagu gecauuucagu gagagaceu guuaaggaa cauuggagag 4200 uauguuqag gecauuucagu gecauuucagu caaquugaaca gecggacuu 4200 uugaagaaga augaagaacu geguucuu caaggagacu caaggagaacu caggagaagg gegggagag geggguccu 4440 uggaugaaga gegauuauug gagagaacucu accaacaacu uggagagag gegggagag geggguccu 4500 uucaagaga guuquagagau gagacuucu geceacaacu geceacaacu aaaaaauuug 4620 uucaagagu guuquagaga guuqu	gcggcucuca	ucgaagaggg	gcagcggaua	gccgagaugu	ugaaguccaa	gauccaaggc	3900
cucgeagaa uqucaecacu gacagcagua cacagagaac cacagagaac quuquagaa 4080 gacgaccuca cacaccacaca gagacacaca guuuquaga 4200 quaggagaa cacaacacaca gagaacaca guuuquaga 4200 quagaagaa qaauqaaca 4200 4200 quagaagaa qaauqaaca 4200 4200 qagaacaca qaauqaaga cacaacacaca qagaacaca 4200 qagaacaca gagaucaaca cacaacacaca qagaacaca 4400 qagaacaca gagaucagaa cacaacacaca qagagaacac 4500 qagaacaca gagaucagaa acagagaaca 4500 qagaacaca gagaucagaa qagagaaca qagaacaca 4500 quagaagaa qagaacaca qagaacaca qagaacaca 4500 quagaagaa qagaacaca qagaacaca qagaacaca 4620 qaacaagaa qaacaagaaca qaacaacaa qaacaagaaca 4620 qaacaagaaa qaacaagaaca qaacaagaaca qaacaagaaca	uugcugcagc	aggccucuaa	gcaggcccag	gacauacaac	ccgcuaugca	ggcuucaugg	3960
gccpccuca ccaquecqu qucqaccaqu accacaaucq cauuqgucqq cauuqgucqq 4200 ugquqqqqqq ccuqaqqqa cauuqqqqq quuuquqqq quuuqqqqq 4200 ugqqqqqqq cquuqqqqq qquuqqqqq qquuqqqq qquuqqqq 4200 ugqqqqqqq qquuqqqq qquuqqqq qquuqqqq qquuqqqq 4200 uqqqqqqqq qqqqqqqq qquuqqqq qqqqqqqq 4200 uqqqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq 4400 uqqqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq 4400 uqqacqqqq qqqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq qqqqqqq 4500 uqqacqqqq qqqqqqqq qqqqqqqq qqqqqqq qqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq qqqqqqq qqqqqqq qqqqqqq qqqqqqq qqqqqqq qqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq qqqqqqqq qqqqqqqqq qqqqqqqqqqqqqqqq qqqqqqqqq qqqqqqqqq qqqq	cccaaagugg	aacaauuuug	ggccagacac	auguggaacu	ucauuagcgg	cauccaauac	
uggiuagegu cccagaucge acaacegeg gggacaceg geuuuguegu caguiggedy 4260 uuuuuggagg gauuuucggg gauuucggg gauuucggg 4260 uuuuuggaga gauuucaaga gauucaucaaga uuuuggaga 4260 uuuuuuggaa gauucaucaa gauucaucaaga uuuuguuca aagagagad 4260 gagaucauca geguuauuge cuuuuguuca aagagaaace aaguuggaca 4400 uggaaggagu gaagaguuca caauuggaaa acuuggueucu 4500 agcuaauua gaagaguucca aacuuggau acuuggueucu aacuuguud 4600 uuuuggueu gaacuugguu gacacuuuu aacuuggaad acuuguguu gacacacuu aagaacaca auguuggggad 4700 uuuuggaga gagacuuuu aauuuggaga gagacuaaa geecacacuu aagaacacac 4800 aacuuggga gaacacacac auuugaacac gaaguugag gaaguugag gaauucaagag gaaauucaagag qaagaacaca aagaguuaca gaaguugag 4900 caaaaacacaa <td< td=""><td>cucgcaggau</td><td>ugucaacacu</td><td>gccagggaac</td><td>cccgcggugg</td><td>cuuccaugau</td><td>ggcauucagu</td><td></td></td<>	cucgcaggau	ugucaacacu	gccagggaac	cccgcggugg	cuuccaugau	ggcauucagu	
gugggggug ccuuaggaca cauaggcug guaaggug uguuaggaag 4260 uuuuguugaag uuuugaaca uuuugaaca ccuuaggaag cuuugagaag ccuugagaag 4300 uuuugaaca geduuuuga uuuugauuca cuuugaaca acaguaggaa ceuugaggag ceuugaggag ceuugaggag 4500 uuugaacaa gaagaacaa acauugaaa acaguaggaa ceccaacaa 4500 ugacagaau gaagacucaa acauugaaa accaacaaca uuugacacaaca daccuacaacaacaa daccaacaacaaacaacaacaaaaaaaggag 4500 uucacuguuca aaaguugaca gaccuacaacaacaacaacaaaaaaaggag 4600 aaacaacaacaacaacaaaaaaaaaaaaaaaaaaaaa	gccgcccuca	ccaguccguu	gucgaccagu	accaccaucc	uucucaacau	caugggaggc	4140
uauggugegg geauuuuuggggg geauuuuggggg ucuuuuggagg quuuuuggggg 4320 uggaguduuuu qoggogcau ucuuuggaga ucuuuggaga 4320 uggagugaaca qoguuguug cacquqqqac cacquqqqac cacquqqqac cacquqqqac cacquqqac qacquuuu deccacauca augucucca augucucca 4500 uacaagggqu qogacququy gacacauca aucuuquucc augucucca augucucca augucucca augucucca augucucca augucucca augucucca augucucca augucucqa qacacauca qacacauca cacaaqacuu caaaaaaa caaaaaaaa qacaqaaaa cuucaqaaaa qaqaaacac qacacaacaa qaaaaaaaa qaaaaaaaa qaaaaaaaa qaaaaaaaa qaaaaaaaa qaaaaaaaa qaaaaaaaa qaaaaaaaaa qaaaaaaaaa qaaaaaaaaa qaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	ugguuagcgu	cccagaucgc	accacccgcg	ggggccaccg	gcuuugucgu	caguggccug	4200
ucuauggaag augucaucaa ucuacugceu gggauccugu cuugguggaag cuugguggaag cuugguggaag cuugguggaag cuugguggaag cuugguagaag cuugguagaag gegeguccaa 4840 uuggaugaagaa gegeuuauug cuuuugug cuuuugug cuuuuggaagaagaagaagaagaagaagaagaagaagaaga	gugggggcug	ccgugggcag	cauaggccug	gguaaggugc	ugguggacau	ccuggcagga	4260
ggggucaucu gcgggccau ucugcgcgc cacgugggac cgggggaggg gcggucaca 4440 uggaugagaca gcguaugacaca gcguaugacaca caccaacuac uuggcucqccc uacucacacac 4560 ugcuacucuca gaagacucca caauuggau gcaccaacuac uugacagacu 4680 uccugacucuca gaagacucca caauuggau gcaccaacuac uucacacauca 4680 uacacagggug guquggccgg cacagcuggguu ggaccacacca gcugcccuu gaggacaca 4680 accuggcagg auuucucgaccu ggacuuuu gagacacag gcugcggac 4680 accugacagg auuucucgacca aauucucgaa 4860 accugacagg caaaacaga gugacagga gccaguacac 4920 cacacacaca aacagagau cuucacacaca aagagauacag gugacaggag gcagaguaca 4920 cacacacaca aagagauac aagacguauu gcagaguaca gaagauacag gcuucagua gaagauacag gaagauacag gaagauacag gaagauacag gaagagaaca gaagagaaca gaagagaaca	uauggugcgg	gcauuucggg	ggcccucguc	gcauucaaga	ucaugucugg	cgagaagccc	4320
uggaugacaca ggcuuauugc cuuugcuuce agaggaacc acgucgccc uacuuacaca 4500 gugacgagau gagaugcguc gaagguug gaagguug gaccaacuac uuuggcuccc adcuuggaac acuuggaac decauuggaac acuuggaacac acuuggaacac acuuggaacacc acuuggaacac decauuggaacacc acuugaacaccc acuugaacacacc acuugaacacaccacacacacacacacacacacacacaca	ucuauggaag	augucaucaa	ucuacugccu	gggauccugu	cuccgggagc	ccugguggug	4380
gugacgagag cgaaugcguc gcaaggugug accaacuac uuugguccucu uacuaacac 4560 agceuacuca agaagacuca caauuggaua accuaacucu uaccaacucu auguccagga 4560 uccugaccucua agaagacuca gcaccaacucu uacaaagggg 4740 uacaagggug uguggccgg cacuugacaccaccu uacacaggug cacuugaga 4800 accuuggaag agaaccuuuc uacacagagag gcacquaggu gcacquagga 2gugccaacc 4800 accuuggaag agaccuuucc uacacagagag gcagagagag gcagauacca 4800 cacaacacacu accagaaguu cacaacagau cagaauacaca agacgaguag gcagauacac 4800 cacaacacacu agacgaguu cucaacaaca cucaacaaca agacgaguu caaacagau cagagagag caaacagau cagagagaccu agaccaaca cacagaaaca caagaaacaa cacagaaaa cacagaacaaca cacagaacaaca caaagaacaacaaca cacagaacaacaacaacaacaacaacaacaacaacaacaa	ggggucaucu	gcgcggccau	ucugcgccgc	cacgugggac	cgggggaggg	cgcgguccaa	4440
agocuacuca gagacuca caaquggau acugagacu gocceaucc augucegga 4620 uccuggcucc goqacugugu ggacuggguu ugacacaucu ugacagacuu caaaaagagg 4740 uacaagggug ugugggcogg cacugacuu augucecca ugugacacaucu ugacagacuu caaaagaggg 4740 uacaaggggu ugugggcogg cacugagacu augucegga gagcuuccu augucaauug ugacacaug ggaccuacaa cugcagaacca 4800 aucucuggga gagacuucc augucaauug ugacacaug gccuacaaca gagagacg caacuggagg gccaugagg ggaggugacg 4880 caacaagagg gagacuuccu uguacacug guggacgggg gccagaacc 4920 ccaacaacaa agccguuucu uguaacagg cugacacaug ggaggugacg ggaggugacg 2980 caacaacaa agccguuuuc cagggauga gucuccuu ggagacacau ggagacacau ggagugagggu caacaacaa agcaguuucu cagggaugag gucuccuu ggagacacau gagacacau acaaggugggu uuucucuuggagg gucucaaaaagagg gagacagaa auuccuucu cucaggaaga gagaccacaa aggagagaca caaggagaga gagacaacaa cuugaagag gagacaacaa cuugaagag gagacaacaa cuugaagag gagacaacaa cuugaagag gacaggagaga auggagagaga aggagagaga gagagaga	uggaugaaca	ggcuuauugc	cuuugcuucc	agaggaaacc	acgucgcccc	uacucacuac	4500
uceuggeucegegacgugugggacuggguugacceucuugacagacuucaaaaauug4680cugaccucuaaauuguucecaagcugeeegucuceuceuucaueucuugqagagaca4800aucucuggaaaugucegeeugggeuuauaugaceaegguugeceuuggecuaaaa4860accuggaaguggaacuuceuaugacaaggaggugegegeeugecagaace4860accuagaaguggaacuuceuaugacaaggaggugegegeeugecagaace4860caacuacacuucuceagaacuuuucucuuggugagaggaggacaauucu4980caacaacacaaaccuucacagauuuuucucuuggugagaggaggacaauucu3040caacaacacaaacceguuuuucaacaagaaggguuucuuuugaguaggaggaguuceau5100cucaacacacaaccaguaacacaaggaaggagacuugaggagacgguuuug2220cucaacacacaaccagaacaccauagaagaacgguuuugacggaguaa5220ggegguguggcacagaacaccauagagagaugacaacacugacauugagacguuugagga520ggegguguugcacagaacaccauagagagaugacaacacugacauugagacguuugagga5400ugacagaggagagagagagagagagagagagagagagaga	gugacggagu	cggaugcguc	gcagcgugug	acccaacuac	uuggcucucu	uacuauaacc	
cugaceucuaaauuguuececaaguugeeagecuceceuuaaaaaggguuaaaaagggu4740uacaaaggguuguugugeegeqaauuguaaggacucaauuaggacacaacgguugeceuau4800accuugeagggagacuuuucuauaaauuguacaaggagggucacuugaace4800accuuggaaggagacuuucuuauaaauuguacaacaggaggucacuugaace4920caacaagagucaacaucucuuuuucucuuggugacacauggucagauacegagauucau5040caacaacacaucuccaacaaaacacuucuucoggaagaagguugacaggugucagauacau5100cccacacaaaageguuucaggagaagguugacggaggucagagaagguugacguugaguucaug5280ccucaucuuaggagacaacucaguagacaagacaggaaggucacaacaagacggagacaa5280ccucaucuuaggagacaaacuuaagacguaagacaggacaugcacaacagaggaccaag5280gagggggggaagaagaggacuuaagacaaaagacaggacuacaaaaaaagagagcaaaauucaaaaaagagagcaaa520aaugacgagagauaaacaaaaaaagaagagaaauaacaaaaagagagcaaaauuacaaaaagagagcaaaauuacaaaaagagagagaaaaaaccaagaaaaaagaacaagaacaaaccaagaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	agccuacuca	gaagacucca	caauuggaua	acugaggacu	gccccauccc	augcuccgga	
uaeaagggugugugggeeggcacuggaucaugaccaegegeugeccuugcggegeaac4800aucucuggcaaaugucegeeugggeucuauugaugacacaeggeecuaaace4860accagaagggaacuuuceuacaauugguacaauugggeecuaaace4920cecacagaagucacacacagaguuuucacauuggguggeggeeueggaguacggagaguacggagaguacg4980cagcaaugggucguacucauuuucacaugguuuucuccugggugagegguugegaguacaaauuccuug5040cecacacaaageeguuuuceggaagaggucucguucugegaugaguaauuccuu5160geugucgggucecageuucceugugaacegagceeageagagceeaugeggueug2220cuaacaagaucecacacaacacageggagacacaugeegagaceguugagaguccau3220geeggaguucecacacaacacaguaggaacacaugeegagaceguuggagaguccau5220geeggguuucucagaacaageeugagueagacagugeaugecaaceugeeugagueegacaugueggacagugeugeeugaguee520ageggguuucacagaggagceuugageeucaaaacaaagaceggeeuugeeugagueegeeugagueeguuugageuguucaaagagaceggaguegeeugagueeguucaaaagageeugaguee <t< td=""><td>uccuggcucc</td><td>gcgacgugug</td><td>ggacuggguu</td><td>ugcaccaucu</td><td>ugacagacuu</td><td>caaaaauugg</td><td></td></t<>	uccuggcucc	gcgacgugug	ggacuggguu	ugcaccaucu	ugacagacuu	caaaaauugg	
aucucuggca auguccgceu gggcucuaug aggaucacag ggccuaaaac cugcaugace 4860 accuggcagg ggaccuuccu uucaauugc uucacaggag gccauugcgc gccgaaaccc 4920 cccacagaacu acaagaccgc caucuggagg guggcggcu cggaguacgc ggagugcgc 4980 caacuaccau cuccagaguu uucuccugg guggacgguu gcagaucac uagguuugc 5040 caacuaccaa agccguuuuu ccgggaugag gucucguuuu gcguugggcu uaauuccuuuu 5160 gcugucgggu cccacaccaa agccguuuuu ccgggaugag gucucguucu gcguugggcu uaauuccuuu 5160 gcugucgggu ccaagcacac cugggagag gucucguucu gcguugggcu uaauuccuuu 5160 gcugucgggu cccacacca caagcaacac cuaugaggg acaggaggac aggggcgag acgggggggggg	cugaccucua	aauuguuccc	caagcugccc	ggccuccccu	ucaucucuug	ucaaaagggg	
accuggóagg ggacculucc laucaauugg lalacaggag gccaguggg gcgaaaccc 4920 cccacgaacu acaagaccg caucuggagg guugcggccu cggaguaggc gagugugg 4980 cagcaugggu cuucculuu uuuuccugg guugaacugu luucuccuug guugaacuga lugacaacug acaaucugaa auuucculug 5040 ccaacaacaa agccguuuuu ccgggaugag cugagacacug acaaucugaa auuucculug 5040 guuguggggu ccaacaac acgcguuuuu cagggaugag guugaaggag gagagagag gagaguacuu gagguuccau 5160 guuguagggu ccaagaaca cacggcggag acuggcgacg gagaguuuu gagguuccau 5220 cuuaacagau aggagacca cacggagaga acugcggcgg ggcguugg aggguccaug 5280 gagggguggg aagagagag caagaacaa gacguuuuu gagaacug aagagagaa aagagagaa caucaagaaca gacguuuuu gagagaccaacaa agaagagaa aagaagagaa cacggaguuuu gagagacca gagguuuu caagacaaga	uacaagggug	ugugggccgg	cacuggcauc	augaccacgc	gcugcccuug	cggcgccaac	
cccacqaacudaaagaccgcaucuuggaggcuucuggaggcuugacggccucggaguacg4980cagcauagggucguacuccuauguacacaggacuugacacacugaaauccuuga5040cccaacacaaagceguuuucegaguugagguucucguucgegaguucauaaguucacuug5100cccaacacaaagceguuuucegagaugagguucucguucgeguugggcuaaguucau5160cuucaacagaucccacacaucacagcaggacuucagugagacagcacacaggeguuugggaggggcauu5280cuucacuucuagagagacacacuuagacguggacauguccagcacacguugucgeggggcacac5340gagggaguuucacaggagaacacacuuagacguggacauguccagaaceguugggg500gagggguuucacaggagacuuagacagacacaugagac4960gagggguuucacaggaccuacacggcuuggecaggacuuggecaggaucacaugecaggaguacacaugaauguaggaggaaggagcagaacacgguugggecacggcuuacacgacacacagecaguugagacagcacacacacacacacacacacacacacacacacacaca	aucucuggca	auguccgccu	gggcucuaug	aggaucacag	ggccuaaaac	cugcaugaac	4860
cagcaugggucguauuccuauguaacaggacugaccacuaacaaucugaaaauuccuug5040caacuaccuucuccagaguuuuuucuccuggguugaacgguguagauucgaggugbiooguguaggggucuccagcaucaagcguuuucuggaagagguugaacgguguagauuggac5100guguacggguceacaccacaagcgguuuccugggaagagguucucguucugeguugggcuaaggggacaa5220cuaacaagauaggegacuccucaguagagaagguagcagaagggagcaac5280gaggggugggcucagacaacacuaugagagaaggugcagagggggccaac5400gagggggugggcucagacaacacuuugagceagggugccagguucagacac5460agaugggggucacaggacaacacucagacacgcucaauacaaacacagcacaccacaggagaca5760aaaugaggagcaccgaccacaagaccacaccacaggaggacacacaggagacacacaggagacacaggaggagaca5760aaauagaacgagucugacaacacacaaacacaagaagaggacacaggagaacacacacacacacacacacacacacacac	accuggcagg	ggaccuuucc	uaucaauugc	uacacggagg	gccagugcgc	gccgaaaccc	
caacuaecuucuccaagaguuuuucuccuggguggacggugugcagauccauagguuugca5100ccacacacaaagccguuucuccugugaaccugucucguucugcguuugggcuuaauuccauu5160cucacacacaagcccacacucaugugaaccugagcccacacgagcugggcuacagguauca5220cucacaucuugaggcgacacaccaaguaacaugagcccacacgagcuuugcacagggacac5280ugacacacaccaagaacaacucaugaagugacauuguucggagcguuugcacgggacacac5340gagggguuugcaaggagacaacucuaugaccuaggagugccaugacaugguucggacauggucggacauggucggaaucgugugcaaggagccaucucuaacacaggacauggucguucuaaacacagacaggucuugacacguugggaaucgugagcacaggacaacuccecaagagaacaggccauuaccagacacacgacaggacaugaccagcuugggacagguucacacaauacaagagggccagauuaccaacacgcacagagaccagacagguccagaggggggac570accauauacaagacccucaagagacaagguuccgacaagguuccgacagguaccacagguagcacacagguagcaca570gggagaacacacaccauagaagaccagguuccgacaagguuccgaccaggacacaccauagaag570gggagaacacacaccauagaagaccagguuccgacaagguuccgaccaggacacaccauaacaaacuccacacaaacuccacacacacacacacacacacacacacacacaca	cccacgaacu	acaagaccgc	caucuggagg	guggcggccu	cggaguacgc	ggaggugacg	
cccacaccaaagecguunuccgggaugaggucucguucugeguugggcuuaauuccuu5160guuguggggucccaqacuucucacagcuuggaggagccegacgcagacguuuugaguccaug5220cuacacaccagecggagacucucagugagccacagcuuuugagguccaug5280geggguguggcucagagacaacucagugagccacaguuuuaagcacaguuggagcacaguuggag5400gugggcguggcucagagaagagaccuggccuuagagguccaa5460auggcgguuuccacaggacaacuuuaagacgugacaggacuuacuucagaacacgcuccaaga5580gaaucguggagagggccagacuuaacaacaccacaggacuuacuucaacacacggagugugacaccacguugg5580gaaucguggagacagcecacacuccacaagagacaggcecagacuucaagaccucaguugguuuuucccacec5580gauaucgugagcuccaagacacacacacacagagaaccacacacacacac	cagcaugggu	cguacuccua	uguaacagga	cugaccacug	acaaucugaa	aauuccuugc	5040
gcugucggguccaagcuucccugugaaccugagcecgaegeagacguauugagguccau5220cuaacagauccgccceacaucucagugagaacuggaggaacugaggagaacugaggaguucgagcguuugacggggacuu5280gcggguguggcucagacaaacucaugaagugaacauggucgaugccaaccugcugaguccugucacauggag340ggggguguggcucagacaaagccuugaguccauggugccggauucuggacuuuucuggacuugcucauggag5400gaggggugaggaggaggaggaggaggaggaggaggaggag	caacuaccuu	cuccagaguu	uuucuccugg	guggacggug	ugcagaucca	uagguuugca	
cuaacagauccgcccacaucacggcggagacugcggcgggcgcuuggacggggacaccuucaucuqaggggaquucucaquagaccacguuquagaccacguuquagaccacguuquagaccacguuquagaccacguuquagaccacguuquagac5400ggcgguguggcucaqaacacacuuqaagucauggacuuqauggacuuqcacguuquagacs400auggcgaguuucucaqaacacacuuqaagucaugguuqcccquucuuqaacuuuucucaqaacac5460auggcggguuucacgggcuuaceggcuuqgacaggccuqacuuacaacaccggagugcaugcuccccagg5520agauuquagaggaggccagauuaccaacacccacacguuqcugguuquagcuccccaag5700accauaucaggcucguccacggagggeggccaccacguuqcugguuguucgacguccccu5820ggugaagcagccccucaagagacacuggccaucaagaccuuuggccagcccccucaga5700accauaucaggcucguccacgggggggggcgcccccaauuuggccagcgacguucccu5820ggugaagcagccccucaagagacacuggccaucaagaccuuuggccagcgacguucccu5820ggugaagcaggcucguuccaaggagggggggggggggggggggggggggggggggggg	cccacaccaa	agccguuuuu	ccgggaugag	gucucguucu	gcguugggcu	uaauuccuau	
ccuccaucugaggegagcuccucaqugagcagcuaucagcaccguucgugeaggcac5340ggcaguaucgacagaacaccuaugaeguggacauggucgaugccaaccugcuuauggag5400ggcgguguggcucagacaaggccugaguccauggugcceguuucuuggacca5460auggcgaggaagagaggaccuuagagccuucaauaacauaggagugcaugcucauggaca5520ageggguuuccacggccuuaceggcuuggcacaggcuuacuacaacacgcegcuugug5580gaaucguggaceccaagaagceccecaagaagacacggcagacagugugugucuucuccecacc5640gagaaucaaggcucguccaaggaggeggggcacacgcgacaguguguguuucuccecacg5760ggugaagcagggcucguccaaggaggeggggcegcegaauceggggucugagggagag5760ggugaagcagccccucaagagucugaucaaggcacgccegaaucegggguccgacgguccagacegggggag5880cuuggaagagaacegguucgggcucggagguuuugucuacuugucuccaccceaggggggggagguucaccceaggggggggagguucaccceagggggag5940ggagagagaaguucgaucaggcucgagguuugcuccaugaaggcuccauauugcuccacagaceagggggag5760gaagaggagaaguuucgauuauacuccugggceuccauuugcuccacacuceagaggggag5940guguacuguuguuccaugaaucaccuugagcucgagauuugcuccacacaceagagggacuceagagggacuceagagggacuguguacuguucaacaucaaagaggcccaagaggaccaagaacaucaacaaacaucaacaaacaucaaca624	gcugucgggu	cccagcuucc	cugugaaccu	gagcccgacg	cagacguauu	gagguccaug	
ugcaccacceacagcaacaccuaugacguggacauggucgaugccaaccugcucauggag5400ggcggguguggcucagacagagccugaguccagggugcccguucuggacuuucucagacca5460auggcggguuuccacgggcuuaccggeuugggaauacgccucaauaccaucggagugacaugcuceccaagg5520gaaucguggggaaggggccagauuaccaaccgcccaccguugcuguugugucucucecccc5640cccaagaaaggccccaacgcuccccaaaggagacgccggacaguugggucugagcgagaggc5700accauaucagagcccuccaggggggggggcccccaaauuuggccagccccccuagag5760ggugagagacggcccccucaagagacagguuccgccuccucuuugcccccccucgaggggggg5880ccuggaguaccgcccguucagggucugaucaggucucaucuuugcuccaccucaaccuccccccaagggggac5940gggguagcucgucucaauguaucacuccuggaccuccucuuaucucagggggacgaggagaa5940gggguagcucgucucaauguaucacuccuggaucacuccuggaccuccaggggcuuauuacaccccuguaggacccccuguaggccc6060gaagaggaaaaguuccaaugucaaccucuugaguacucccaaaagguacccuguaggccccuguagaccccauaaacaac6120aaggucaggaguccaagggccuaagaggccuaaaagguaccaguugacaccaguugacaccaguugacacccaguugacacccaccuaaacaca6120aaggucagagcaagaguccaaguugacaccaguugacacccaguugacacccaccuaacaccaaauuacaccaacacacacacacacacacacacacacaca	cuaacagauc	cgccccacau	cacggcggag	acugcggcgc	ggcgcuuggc	acggggauca	5280
ggcggugugugcucagacagagccugaguccagggugcceguucuggacuuucucugagcuu5520auggcegguuucacgggcuuaccuugagcecuucuugagcuugcuceccaag5580gauucguggaggaggccagauuaccaacegcccaacggugcuugguuguguucecccacg5580cacauaucagggaggccagauuaccaacegcccaacaguugcuuguuguguucecccec5640cccaaagaaggccccgacgccuuceccaaagagacgacggacaguuggucugaaggagaga5700ggugagccggccccucaaggggggggggcgcegccgaaucaguggggggggacgucccc5820ggugagccggccccucaaggucuguuccaggucuguuccaggucuguuccagguagagcuucaccucccccccagggggggg580cugggagagaaccggaccuggagucuguuccagguugucucuuuaccucccccccagggggggg5940gggguagcuccccguuugggcucgggguucaacgacggguuaacgacggguuaacgacgagaa6000gaagaggaaaaguugccaaucaacacuuuuaguaacucguaaaaggacaaaaaaggacaa6120guuucuguucaacacuuuaguaacucguaaaaggagacaauuuugacuuaaaaggagaaacuugacgaacuugacga6120aaggucagcgcaagguuccucaagggccuaguuguacugaguuuuacaggaaaaggagaaaagcggcuuc6240aaggucagcucaagaauacaggaggacaaggagggccaaggguuguacagagaccacaaaaacacaauucccaaca6420aacacacacacaaacaucaagaagguguucugguaggagaccucacaaaaaacaaauucccaaca6420gacauuacaa </td <td>ccuccaucug</td> <td>aggcgagcuc</td> <td>cucagugagc</td> <td>cagcuaucag</td> <td>caccgucgcu</td> <td>gcgggccacc</td> <td></td>	ccuccaucug	aggcgagcuc	cucagugagc	cagcuaucag	caccgucgcu	gcgggccacc	
auggecgaggaagagagcgaceuugageeeucaauaccaueggagugcaugeueeeeag5520ageggguuuegaaggecagauuaacaaaeegeeaggeeugggeaeggeeugaeuacaaeeegeegeuegg5580cecaagaaggeecegaceeuuaecaaaeegeeaaeggeeugeuuguugugeueeeeeeee5640gugaugeageecegaceeuuaecaaaegeeaaeggeegaeaguggggeeeagugggggee5700gugaageeggeeceeucagaggaggggggegeegeegaaueegggggueeeeggggggg5760gugaageeggeeceeucagagaeaggueegeegeegaaueeggggggg5880eeggguageeeegguueggggueugaueaggueugaueaggueugagageueegagggggg5940gaggagagaaeegguueggeeceeuugggueugaueaugeueegagaaeegagggggg5940gaggagagaaeecegguueggeecegggueeuugueuaeuugeueegagaaeeceggggeeeeceggggeeeeceggggee6000gaagaggaaaaguugeeaaueaaeeecuuugaguaaaeuegeuuuuagaaeueeeceeacuugaggaaaaggueaeeeceeacuuggaggaggeegueeaaaagguaaeeeceeacuuuu6120gaagaeageeeaagaeueeeaaeeuueaggaaggeeueegagggeeueaaaaggueaeeeceeauuuu6300gaeagaaueeeacaaaaaaggaguueueaggagggeeaaggeuugaaeeegeuugaaeeeeeceeauuuu6300gaeauuaeaeeacauuaeaegaguueueeguugaaeeeegeuugaaeeeeecaaaaaee6420aacauuaeaeaaaaaaeeegaguggaaeueecaaaggagaeecaaaggagaa	ugcaccaccc	acagcaacac	cuaugacgug	gacauggucg	augccaaccu	gcucauggag	5400
ageggguuuccacagggccuuaccaggcuuuggcacaggccuuacuacaacccgcacagguuuguugtouccccace5640cccaagaaggcccaagaaggcccaaagaaggcccaacaguugcuugguugugucucccccc5640accauaucagaagcccuccagcaacuuggcaagacgccggacaguuggucugagcagagac5700accauaucagaagcccuccaggaggaggggccgcgaauccgggguccgaccaugagg5700ggugaugcagccccucucaagacaagguucgccuccucuauugccccccucccucgaggggg5820ggugaagcagccccucucaagacagguucgcuccucuuugccccccuccaaggggggg5880gggguagcucccgguucgggcucggggucuuguucuacuugcuccaugaggaggacgagaac6000accgugaggaagguuccauuaaccacucuuaaccacucauaaccacacacaccaagaagaa6120guguacuguuaaguugccaaaaguugccaaaacaccuuugaaaacacacaagaagaaaacaacaagaagaaauuuugacaag6180aaguuagaccaaaguuggauucaagagaccucuugaagagcaaauuugaacaaaauuugaacaaaaguugccaaaguugaccaaaguugaccaaaauuucgaagaggaggaggaggaggaggaggaggaggaggaggag	ggcggugugg	cucagacaga	gccugagucc	agggugcccg	uucuggacuu	ucucgagcca	
gaaucgugga ggaggccaga uuaccaaccg cccaecguug cugguugugc ucucccccc 5640 cccaagaagg ccccgacgc uccccaagg agacgccgga caguuggguu gagcgagagc 5700 accauaucag aagcccucca gcaacuggcc aucaagaccu uuggccagcc ccccucgagc 5760 ggugaugcag gcucgucaca ggagguggcggg ggcgcgcggaau ccggggguc gacggucccc 5820 ggugagccgg ccccucaga ggaggucuc gccucucua uuggccagcc ccccucgagc 5760 ccuggagauc cggaccucga gacagguucc gccucucua ucggcgguc gacggucccc 5820 ggugagccgg ccccucaga gacagguucc gccucucua ucggcgguc gacgguggggggg 5940 gggguagcuc cggguucggg cucggggucu uagucacuu gcuccaagag gacggagag 5940 gggguagcuc cgggucgg cucggggucu uagucacuu gcuccaagag gacggagag 5940 gggguagcuc cggguucga auacuccugg aguaacucgc caagggggaa aguugccaau aacaccuuug aguaacucgc uuaugacagg gacgaauacc 6000 aacgguagcuucu caaccuuuug aguaacucgc uuaugacagg gaggacgauac aaguugacaga gaggcccaa uuaugacuca aaguucagaagacuc caagagggcua aaaagguaac uuuuugacagg gagggccgcuc caagggggca gaggaccuc aaguugacagg gagggccguc aaggacccc aaguugacagg gagggccguc caaggagacc caaaaaaaa gguguucugc guggacccc gcaaagaggg gagggccguu aagaaacac 6480 gcucgccuaa aaaagcuucc cucgguggagaa aaaaggagac ccaaaaaacac aaguucgaaaaccaauaaga gguguucugc gucgggucu gcaggagaaaa cuccaaaaacac gggggggggg	auggccgagg	aagagagcga	ccuugagccc	ucaauaccau	cggagugcau	gcuccccagg	
cccaagaagg ccccgacgc uccccaagg agacgccgga cagugggucu gagcgagag 5700 accauaucag aagcccucca ggaggcggc gacacuggac gccgccgaau ucggcccccu caguggaguc gguguggcgg cccccucaga gacagguucc gccuccucua ugccccccu caggggggg 5760 ggugagccgg cccccucaga gacagguucc gccuccucua ugccccccu caggggggg 5880 ccuggaguac cggaccugga gucugaucag gucugaucau ugccccccu caggggggg 5940 gggguagcuc ccgguucggg cucggggucu aaccucuca accuuccac cuguagcac cuggagguac cucggggucu aucucuug aguaacuccu cuguagcac cugugugaacugu caaccuuuug aguaacucgu uuuuugacaau caaccucuuug aguaacucgu uuuuugacaag 6120 guguacugua caacaucaaa gagcgccuca cagaggggua aacaacaaaa aguucgcaau caaccuuuugagg gagggcggcaaggucaaaaaaaaga acaacaacaaa aguacggaacaa gagguccaaa gagguccgaa gagguccgaa accacacaaca aguacgguuc caacaucaagg gagguccgaa gagguccgaa accacacaaca aguacgguuc gaaggaccu cuggaagacc cacaaacacac aauccaaaaga guguuucug gugagaccac cacaaacacac aauccaaaaga gagguucug gugagaccac cacaaacacac aaucccaca 6420 gacauuacac aaaagcuucc ucaggcggua aacggggggg gaggagaaaaa gggcccuuau cacacacacacacacacacaaaaga gaguucuug gugagaccac cacaaacacac aauuccacac 6420 gacauuacac aaaagcuucc ucaggcggua aacggggggg gaggagaaaca ggggagaaaaa gggagaacacacac	agcggguuuc	cacgggccuu	accggcuugg	gcacggccug	acuacaaccc	gccgcucgug	
accauaucag aagccucca ggaggggggg guggagacggg cccccucagag gugaggcggg cccccucagag gucaggagggggggggg							
ggugaugcag ccccucaga gacaggucc gccuccucua ugccccccu gaggggggg 5880 ccuggagauc cggaccugga gucugaucag guagagcuu aaccucccc caggggggg 5940 gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000 accgugugu gcuccauguc auacuccugg accggggcuc uaauaacucc cuguagccc 6060 gaaggaggaa aguugcaaaa gagggccuca caacccuuug aguacucgc uguugcgaua ccauacaaa gaggcgcuca aagggucaggu acaacgaagugc aaggucagc caacgcuucu caccuugagg gagggguaguucagguucgaaagguaaguc aaggucagg aaggacgaagugc aaggucaga gaggacgaaggac aacacaacaa gagggcgaaggaag	cccaagaagg	ccccgacgcc	ucccccaagg	agacgccgga	cagugggucu	gagcgagagc	
ggugageegg cececueaga gacagguuce geueeueua ugeeeeeeu egaggggag 5880 eeuggagaue eggaceugga gueugaucag guagageuue aaceueeee ceagggggg 5940 gggguageue eegguueggg euegggguu uggucuaeuu geueegagga ggacgauaee 6000 aeegugugeu geueeauu eaaceeuuug aceggggee uaauaaeuee euguageeee 6060 aeegugugeu eaacaueaa gageeeuea eaaggggeua aaaagguaae uuuugacaag 6120 uuaugacagg eaaggueeea uuaugaeuea gueuuaaagg acaucaageu ageggeeuea eaaggueagee aaggueega gaggeeuea aaaagguaae uuuugacagg 6180 aaceaaguee aguauggauu eggggeeaag gaggeeguee aguugaeuee aceecauuuu gagggeegue aaguugaeue aceecauuuu 6300 gaagaacea aguauggauu eggggeeaag gaggueegaa geuugueegg gagggeeguu 6360 aaceacauca agueegugg gaaggaceue euggaagaee eaaaaacae agueegggu gaaggaceue euggaagaee eaaaaacae aaaageuuee uuaageegg guugaeeeeg eeaaaaaae gguuaeee aaaageuuee eeaaaaaee aaageuuee uuaageegg guugagaeee eeaaaaaae gguugaeee eeaaaaaae gguugaeeeg guugaeeeg eeaaaggggg uaaaaaaee 6420 aeaaaaaaa gguugaeee uuaggeggua aaaggagaea eeaageee uuaggagaee uuuaegaagae uuuaegaagae uuuaegaaa eeaggeega gagggeeau uuaagaaaaee eeaaggagaea uuuaegaagae uuuaegaaa agggeeeau uucaacaagea aacaugeaaa aacaugeaa ggugaaeaee eaaggagaea uuuaegaagae uuuaegaaaa uuuaegaaa agggeeeau uucaacaagea ggagaeaaa aacaugeaa aacaugeaa aacaugeaa agagaeeaa aacaggaaee uuaagaaaee eaaggeeeaa aacaageaa ggagaeeaaa aacaugeaaa agaggaeaaa aacaggaaee aacaageaaa aacagaeaaa aacaageaaa aacaageaaaaaaaaaa	accauaucag	aagcccucca	gcaacuggcc	aucaagaccu	uuggccagcc	ccccucgagc	
ccugagaauc cggaccuga gucugaucag guagagcuuc aaccucccc ccagggggg 5940 gggguagcuc ccgguucggg cucggggucu uggucuacuu gcuccgagga ggacgauacc 6000 accgugugcu gcuccauguc auacuccugg accggggcuc uaauaacucc cuguagcccc 6060 gaagaggaaa aguugccaau caacccuuug aguaacucgc uguugcgaua ccauaacaag 6120 guguacugua caacaucaaa gagggccuca aagggggcu uaauaacucc cuguagcccc 6240 aaggucaggg caaggcuccu caccuuggag gaggggugc aguugacucc acccauucu 6300 gcaagaucca aguauggauu cggggccaag gagggccgca gguggaccaacaacaacaacaacaacaacaacaacaacaacaa	ggugaugcag	gcucguccac	gggggcgggc	gccgccgaau	ccggcggucc	gacguccccu	
gggguageue cegguueggg cueggggueu uggueuaeuu geueegagga ggaegauaee 6000 acegugugeu geueeaugue auaeueeugg aceggggeue uaauaaeuee cuguageeee 6060 gaagaggaaa aguugeeaau caaceeuuug aguaaeuege uguugegaua ceauaaeaag 6120 guguaeugua caacaucaaa gagegeeuea cagagggeua aaaagguaae uuuugaeagg 6180 acgeaaguee caaggeueeu caceuuggag gaggeguee aguugaeuee aceeeauueu 6300 geaagaueea aguuegguu ggaaggaeeue cuggaagaee aguugaeuee aceeeauueu 6300 geaagaueea aguuegguu ggaaggaeeue cuggaagaee cacaaaeaee aauueeeaa 6420 aceaueaugg caaaaaauga gguguueuge guugaeeeeg caaagggggg uaagaaaeea 6480 geuegeeuea ueguuuaeee ugaeeuegge gueegggueu geggaaaaau ggeeeueuau 6540 gacauuaeae aaaageuuee ucaggeggua augggageeu eeuagggggg uaagaaaeea 6600 eeugeeeaa ggguggagaa uuueggagggua acauuggagaeu ceaagaaaae 6600 eeugeeeaa ggguggagaa acauegggggg uaaaaagaagaa ceceauggu 6600 uuuueguaug auaeeegaug cuuegaeuea acegueaeug gagagaeeu ceaaggagae cecaaugguu 6600 aagaeeauau aceaggeeug cuceeugeee gaggaggeee gaaaagaaaga ceceaugguu 6780 aagaeguugee gegeeaageg ggugeuaaee acuageaug guaaaaeea cugegguuae 6840 agaeguugee gegeeageg ggugeuaaee acuageaug guaaeaeeau cacaugeuua 6900 gugaaageee uageggeeug caaggeeeaug uueaaeagaa ggggeeeau cacaagea cuceaagae 6960 caggggaeuu aggaggaea geggaaeeug gagaaeeuu caggaeeau cacaagea 7020 ucugeeeeue cuggugauee ceceaageeg gaauaugaee gggaaaeuu aacaueeugu 7080 ucucucaaaug ugueuguge guugggeee gaauaugaee geggaaaeua cacaageaa acaueeugu 7080 ucucucaaaug ugueuguge guugggeee gaauaugaee geggaaaeua cacaageaa cuugaaeaa acuceugu 7080 ucucucaaaug ugueuguge guugggeeee ggaauauuae caggaeeaa cuugaeeaga 7140							
accgugugeu geuccauguc auacuccugg accgggeuc uauacacucc cuguagecee 6060 gaagaggaaa aguugecaau caaccuuuu aguaacucge uguugegaua ccauaacaag 6120 guguacugua caacaucaaa gagegecuca cagagggeua aaaagguaac uuuugacagg 6180 acgcaaguge ucgacgeca uuaugacuca gucuuaaagg acaucaagcu ageggeuuce 6240 aaggucageg caaggeuccu caccuuggag gaggegugee aguugacuce accecauucu 6300 geaagaucca aguauggauu eggggeeaag gaggucegea geuugucegg gagggeeguu 6360 aaccacauca agucegugug gaaggacuc euggaagace cacaaacacc aauuccaca 6420 accaucaugg ccaaaaauga gguguucuge guggacceeg ccaagggggg uaagaaacca 6480 geuegecuca ucguuuacee ugaccucgge gucegggucu gegagaaaau ggeeeucuau 6540 gacauuacac aaaagcuuce ucaggeggua auaggagcuu ecuauggeuu ecaguacuce 6600 ccugeccaac ggguggagua ucucuuugaaa gcaugggegg aaaagaagaa ceccauuggu 6660 uuuucguaug auaccegaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggecug cucecugee gaggaggee geacugeau acacucgcug 6780 acugagagac uuuacguagg agggeecaug uucaacagaa agggucaaac cugegguuac 6840 agacguuge gegecagegg gaugeuaace acuagcaug guaacaccau cacaugcuau 6900 gugaaaagce uageggecug caaggecgg gagauaguug egeccacaau cucagaaage 6960 caggggacuu aggaggacga eeccagacgg gaggaccuu agaggeecau gaccagguac 7020 ucugecccuc cuggugaue eeccagaccg gaauaugace uggagecaau aacauccugu 7080 ucucucaaaaug ugucuguge guugggeecg eggggeegee geagauacua ecugaccaga 7140	ccuggagauc	cggaccugga	gucugaucag	guagagcuuc	aaccuccccc	ccaggggggg	
gaagaggaaa aguugccaau caaccuuug aguaacucgc uguugcgaua ccauaacaag 6120 guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuugacagg 6180 acgcaagugc ucgacgcca uuaugacuca gucuuaaagg acaucaagcu agcggcuucc 6240 aaggucagcg caaggcuccu caccuuggag gagggcgugcc aguugacucc acccauucu 6300 gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccguug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua auugggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac gguggagga ucuccuugaaa gcaugggggg aaaaagaaga ccccauugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug gagagacau caggaccgag 6720 gaguccauau accaggccug caaggcccaug uucaacagca aggguccaaac cugcgguuac 6840 agacguugcc gagccacagg ggugcuaacc acuagcaug guaacaccau cacaugcuau 6900 gugaaaagccc uagcggccug caaggccug gagauaguug cgcccacaau cucagaaagc 6960 caggggaccu cuggugaucc ccccaagaccg gaauaugacc uggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccaagaccg gaauaugacc cggaggccau aacaccau ccugaccaga 7140							
guguacugua caacaucaaa gagcgccuca cagagggcua aaaagguaac uuuuugacagg 6180 acgcaagugc ucgacgcca uuuugacuca gucuuaaagg acaucaagcu agcggcuucc 6240 aaggucageg caaggcuccu caccuuggag gagggcgugcc aguugacucc acccauucu 6300 gcaagauca aguucgguu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccaucau aguccgugug gaaggaccuc cuggaagacc cacaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 gcucgccuca ucguuuaccc ucagcggua augggagcuu ccuauggguu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggggg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agaggagcau caggagcau caggaccau accaucaggu 6720 gaguccauau accaggcgg gaggccaug uucaacagca aggguccaac cugcgguucc 6840 agacguugcc gcgccaggg ggugcuaacc acuaggaggc gaaagagacc cacaugcua 6900 gugaaagccc uagcggcug caaggccgg gggaaacuu cggagccau gaccagguac 6960 caggggacug aggaggaca gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc cccaagaccg gaaaauugacc uggagcaaa acauccugu 7080 uccucaaaug ugucugugc guugggccg guugggccg gcagauacua ccugaccaga 7140							
acgcaagugc ucgacgcca uuaugacuca gucuuaaagg acaucaagcu ageggcuucc 6240 aaggucagcg caagguccu caccuuggag gaggcgugcc aguugacucc acccauucu 6300 gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cucccugcc gaggaggccc gcacugccau acacucgcug 6780 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggccug caaggcugca gggaaacuu agaggccau gaccagguac 7020 ucugcccuc cuggugacc cccaagaccg gaauaugacc uggagccaa acauccugu 7080 ucucucaaaug ugucugugc guugggcccg cggggccgcc gcagauacua ccugaccag 7140							
aaggucagcg caaggcuccu caccuuggag gaggcgugcc aguugacucc accccauucu 6300 gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cucccugcc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggcug caaggcugc gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggagacga gcggaaccug agaggccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucugugc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140	guguacugua	caacaucaaa	gagcgccuca	cagagggcua	aaaagguaac	uuuugacagg	6180
gcaagaucca aguauggauu cggggccaag gagguccgca gcuuguccgg gagggccguu 6360 aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauucccaca 6420 accaucaugg ccaaaaauga gguguucugc guggaccccg ccaagggggg uaagaaacca 6480 gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagaa ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cuccuugcc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggccug caaggcugc gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggagacca gcggaaccug agaggccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau accucugu 7080 uccucaaaug ugucugugc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140	acgcaagugc	ucgacgccca	uuaugacuca	gucuuaaagg	acaucaagcu	agcggcuucc	6240
aaccacauca aguccgugug gaaggaccuc cuggaagacc cacaaacacc aauuccaca 6420 accaucaugg ccaaaaauga gguguucugc guggacccg ccaagggggg uaagaaacca 6480 gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cucccugcc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggcug caaggcugc gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggagcca gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140							6300
accaucaugg ccaaaaauga gguguucugc guggacccg ccaagggggg uaagaaacca 6480 gcucgccuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cuccugccc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggcug caaggcugc gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggaggacga gcggaaccug agaggccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucugugc guugggcccg cggggccgcc gcagauacua ccugaccag 7140						5-555-5-	
gcucgcuca ucguuuaccc ugaccucggc guccgggucu gcgagaaaau ggcccucuau 6540 gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cucccugccc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggcug caaggcugc gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggagacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucugugc guugggcccg cggggccgcc gcagauacua ccugaccag 7140	aaccacauca	aguccgugug	gaaggaccuc	cuggaagacc	cacaaacacc	aauucccaca	6420
gacauuacac aaaagcuucc ucaggcggua augggagcuu ccuauggcuu ccaguacucc 6600 ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cucccugccc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggcug caaggcugc gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggagacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc gcagauacua aacauccugu 7080 uccucaaaug ugucugugcc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140	accaucaugg	ccaaaaauga	gguguucugc	guggaccccg	ccaagggggg	uaagaaacca	6480
ccugcccaac ggguggagua ucucuugaaa gcaugggcgg aaaagaagga ccccaugggu 6660 uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cucccugecc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggccug caaggcugc gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggagacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140	gcucgccuca	ucguuuaccc	ugaccucggc	guccgggucu	gcgagaaaau	ggcccucuau	6540
uuuucguaug auacccgaug cuucgacuca accgucacug agagagacau caggaccgag 6720 gaguccauau accaggccug cucccugccc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggagacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140	gacauuacac	aaaagcuucc	ucaggcggua	augggagcuu	ccuauggcuu	ccaguacucc	6600
gaguccauau accaggccug cucccugcc gaggaggccc gcacugccau acacucgcug 6780 acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggaggacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140							6660
acugagagac uuuacguagg agggcccaug uucaacagca agggucaaac cugcgguuac 6840 agacguugcc gcgccagcgg ggugcuaacc acuagcaugg guaacaccau cacaugcuau 6900 gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggaggacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140							
agacguugce gegecagegg ggugeuaace acuageaugg guaacaceau cacaugeuau 6900 gugaaagee uageggeeug caaggeugeg gggauaguug egeccacaau cucagaaage 6960 caggggacug aggagacga geggaaceug agageeuuca eggaggeeau gaceagguae 7020 ucugeeecuc cuggugauce ecceagaceg gaauaugaee uggageuaau aacauceugu 7080 uceucaaaug ugucugugge guugggeeeg eggggeegee geagauacua ecugaeeaga 7140						, ,	
gugaaagccc uagcggccug caaggcugcg gggauaguug cgcccacaau cucagaaagc 6960 caggggacug aggagacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc ccccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140							
caggggacug aggaggacga gcggaaccug agagccuuca cggaggccau gaccagguac 7020 ucugcccuc cuggugaucc cccagaccg gaauaugacc uggagcuaau aacauccugu 7080 uccucaaaug ugucuguggc guugggcccg cggggccgcc gcagauacua ccugaccaga 7140							
ucugeccue euggugauce ecceagaceg gaauaugace uggageuaau aacauceugu 7080 uceucaaaug ugucugugge guugggeeeg eggggeegee geagauacua eeugaeeaga 7140						, ,	
uccucaaaug ugucugugge guugggeeeg eggggeegee geagauacua eeugaeeaga 7140							
gacccaacca cuccacucge eegggeugee ugggaaacag uuagacacue eecuaucaau /200							
	gacccaacca	cuccacucgc	ccgggcugcc	ugggaaacag	uuagacacuc	cccuaucaau	1200

```
ucauggcugg gaaacaucau ccaguaugcu ccaaccauau ggguucgcau gguccuaaug 7260
acacacuucu ucuccauucu cauqquccaa qacacccuqq accagaaccu caacuuuqaq 7320
auguauggau caguauacuc cgugaauccu uuggaccuuc cagccauaau ugagagguua 7380
cacgggcuug acgccuuuuc uaugcacaca uacucucacc acgaacugac gcggguggcu 7440
ucageceuca gaaaacuugg ggegeeaeee eucagggugu ggaagagueg ggeuegegea 7500
gucagggegu cecucaucuc ceguggaggg aaageggeeg uuugeggeeg auaucucuuc 7560
aauugggegg ugaagaccaa geucaaacuc acuccauuge eggaggegeg ecuaeuggae 7620
uuauccaguu gguucaccgu cggcgccggc gggggcgaca uuuuucacag cgugucgcgc 7680
georgaeece geucauuaeu euueggeeua eueeuaeuuu ueguaggggu aggeeueuue 7740
cuacuccccq cucqquaqaq cqqcacacac uaqquacacu ccauaqcuaa cuquuccuuu 7800
uuucuucccu ucucaucuua uucuacuuuc uuucuuqquq qcuccaucuu aqcccuaquc 7920
acqqcuaqcu quqaaaqquc cquqaqccqc auqacuqcaq aqaquqccqu aacuqqucuc 7980
ucuqcaqauc auqu
<210> 9
<211> 340
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic RNA
<400> 9
accugececu aauaggggeg acaeucegee augaaucaeu eeccugugag gaacuacugu 60
cuucacgcag aaagcgccua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccuccq qqaqaqccau aquqqucuqc qqaaccqquq aquacaccqq aauuqccqqq 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu ugugguacuq ccugauaggg 300
cqcuuqcqaq uqccccqqqa qqucucquaq accquqcacc
<210>. 10
<211> 340
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic RNA
<400> 10
accegeeceu aauaggggeg acaeucegee augaaucaeu eeceugugag gaacuacugu 60
cuucacgcag aaagcgucua gccauggcgu uaguaugagu gucguacagc cuccaggccc 120
cccccucccg ggagagccau aguggucugc ggaaccggug aguacaccgg aauugccggg 180
aagacugggu ccuuucuugg auaaacccac ucuaugcccg gccauuuggg cgugcccccg 240
caagacugcu agccgaguag cguuggguug cgaaaggccu uguqguacug ccugauaggg 300
ugcuugcgag ugccccggga ggucucguag accquqcacc
                                                                340
<210> 11
<211> 236
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic RNA
```

```
<400> 11
ageggeacae acuagguaca cuccauageu aacuguuccu uuuuuuuuuu uuuuuuuuu 60
uauucuacuu ucuuucuugg uggcuccauc uuagcccuag ucacggcuag cugugaaagg 180
uccquqaqcc qcauqacuqc aqaqaquqcc quaacuqquc ucucuqcaqa ucauqu
<210> 12
<211> 232
<212> RNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic RNA
<400> 12
ииииииии иииииииии ииииииии ииисссиси исиисссиис исаисииаии 120
cuacuuucuu ucuuggugge uccaucuuag cccuggucac ggcuagcugu gaaagguccg 180
ugageegeau gaeugeagag agugeeguaa euggueueue ugeagaueau gu
<210> 13
<211> 17
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 13
cgggagagcc atagtgg
                                                           17
<210> 14
<211> 19
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: synthetic DNA
<400> 14
agtaccacaa ggcctttcg
                                                           19
<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 15
ctgcggaacc ggtgagtaca c
                                                           21
```

```
<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: synthetic DNA
<400> 16
aacaagatgg attgcacgca
                                                                    20
<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 17
                                                                    20
cgtcaagaag gcgatagaag
<210> 18
<211> 30
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence: synthetic DNA
<400> 18
gcactctctg cagtcatgcg gctcacggac
                                                                    30
<210> 19
<211> 28
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 19
                                                                    28
cccctgtgag gaactactgt cttcacgc
<210> 20
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
```

```
<400> 20
                                                                    24
ccgggagagc catagtggtc tgcg
<210> 21
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 21
                                                                    30
ccactcaaag aaaaagtgtg acgagctcgc
<210> 22
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 22
                                                                    18
ggcttgggca cggcctga
<210> 23
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 23
                                                                    30
gcggtgaaga ccaagctcaa actcactcca
<210> 24
<211> 21
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: synthetic DNA
<400> 24
                                                                    21
agaacctgcg tgcaatccat c
<210> 25
<211> 23
<212> DNA
<213> Artificial Sequence
```

<220> <223> Description of Artificial Sequence: synthetic DNA	
<400> 25 cccgtcatga gggcgtcggt ggc	23
<210> 26 <211> 27 <212> DNA <213> Artificial Sequence	
<220> <223> Description of Artificial Sequence: synthetic DNA	
<400> 26 accagcaacg gtgggcggtt ggtaatc	27
<210> 27 <211> 18 <212> DNA <213> Artificial Sequence	
<220> <223> Description of Artificial Sequence: synthetic DNA	
<400> 27 ggcacgcgac acgctgtg	18
<210> 28 <211> 30 <212> DNA <213> Artificial Sequence	
<220> <223> Description of Artificial Sequence: synthetic DNA	
<400> 28 agctagccgt gactagggct aagatggagc	30
<210> 29 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> Description of Artificial Sequence:synthetic DNA(primer)	
<400> 29 aacaagatgg attgcacgca	20

```
<210> 30
<211> 20
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:synthetic
      DNA(primer)
<400> 30
                                                                    20
cgtcaagaag gcgatagaag
<210> 31
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:synthetic DNA
                                                                    30
gcactctctg cagtcatgcg gctcacggac
<210> 32
<211> 28
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence:synthetic DNA
<400> 32
                                                                    28
cccctgtgag gaactactgt cttcacgc
<210> 33
<211> 24
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 33
                                                                    24
ccgggagagc catagtggtc tgcg
<210> 34
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
```

```
<400> 34
                                                                     30
ccactcaaag aaaaagtgtg acgagctcgc
<210> 35
<211> 18
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence:synthetic
      DNA (primer)
<400> 35
                                                                     18
ggcttgggca cggcctga
<210> 36
<211> 30
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 36
                                                                     30
gcggtgaaga ccaagctcaa actcactcca
<210> 37
<211> 21
<212> DNA
<213> Artificial Sequence
<223> Description of Artificial Sequence::synthetic DNA
<400> 37
                                                                     21
agaacctgcg tgcaatccat c
<210> 38
<211> 23
<212> DNA
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence::synthetic DNA
<400> 38
                                                                     23
cccgtcatga gggcgtcggt ggc
<210> 39
<211> 27
<212> DNA
<213> Artificial Sequence
```

<220>	
<223> Description of Artificial Sequence::synthetic DNA	
<400> 39	
accagcaacg gtgggcggtt ggtaatc	27
<210> 40	
<211> 18	
<212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence::synthetic DNA	
.100. 10	
<400> 40 qqaacqcqac acqctqtq	18
ggaacgcgac acgccgcg	10
<210> 41	
<211> 30 <212> DNA	
<213> Artificial Sequence	
<220>	
<223> Description of Artificial Sequence::synthetic DNA	
<400> 41	
agctagccgt gactagggct aagatggagc	30